

Инфотека:

ISSN 1450-9687 (штампано издање)

ISSN 2217-9461 (е-издање)

Часопис за дигиталну хуманистику

Београд, год. 16, бр. 1/2, август 2016.

Импресум

ЗА ИЗДАВАЧА:

проф. др Александар Јерков

office@unilib.bg.ac.rs
Универзитетска библиотека „Светозар Марковић“

Филолошки факултет, Универзитет у Београду

ИЗДАВАЧ:

Филолошки факултет, Универзитет у Београду

Универзитетска библиотека „Светозар Марковић“

Заједница библиотека универзитета у Србији

ГЛАВНИ И ОДГОВОРНИ УРЕДНИК:

проф. др Цветана Крстев

cvetana@matf.bg.ac.rs
Филолошки факултет, Катедра за библиотекарство
и информатику

ОПЕРАТИВНИ УРЕДНИК:

др Александра Трговац

aleksandra@unilib.bg.ac.rs

Универзитетска библиотека „Светозар Марковић“

УРЕДНИК ЕЛЕКТРОНСКОГ ИЗДАЊА:

Јелена Андоновски

andonovski@unilib.bg.ac.rs

Универзитетска библиотека „Светозар Марковић“

УРЕЂИВАЧКИ ОДБОР:

проф. Александра Вранеш, проф. Александар Јерков, проф. Биљана Дојчиновић, Филолошки факултет, Универзитет у Београду; проф. Елизабет Бур, Институт за романистику, Универзитет у Лајпцигу; проф. Владан Девеџић, Факултет организационих наука, Универзитет у Београду; др Милена Добрева, Факултет за медијске и когнитивне науке, Универзитет на Малти; др Томаж Ерјавец, Одсек технологије знања, Институт „Јозеф Стефан“; проф. Светлана Коева, Институт за бугарски језик; проф. Денис Морел, проф. Агата Савари, Универзитет Франсоа Рабле из Тура; проф. Иван Обрадовић, Рударско-геолошки факултет, Универзитет у Београду; проф. Гордана Павловић Лажетић, проф. Душко Витас, Математички факултет, Универзитет у Београду; проф. Катерина Здравкова, Факултет за информатичке науке и компјутерско инжењерство, Универзитет „св. Кирил и Методије“ у Скопљу

ИЗРАДА ВЕБ ПОРТАЛА И ЧАСОПИСА:

Вукосав Средојевић, Универзитетска библиотека „Светозар Марковић“

ЛЕКТОРИ ЗА ЕНГЛЕСКИ ЈЕЗИК ЈЕЗИК::

Јелена Митровић, Филолошки факултет, Београд

др Зоран Ристовић, Основна школа „Мито Игумановић“, Косјерић

ЛЕКТОР ЗА СРПСКИ ЈЕЗИК:

др Александра Павловић, Универзитетска библиотека „Светозар Марковић“

ДИЗАЈН И ПРИПРЕМА ЗА ШТАМПУ:

Александар Милошевић, Бранислава Шандрих и Инфотека тим

РЕДАКТОР ВИБЛИОГРАФИЈЕ И УДК КЛАСИФИКАЦИЈА:

Наташа Дакић, Универзитетска библиотека „Светозар Марковић“

РЕДАКТОР ДОИ БРОЈЕВА:

др Милош Утвић, Филолошки факултет, Београд

РЕДАКЦИЈА ЧАСОПИСА:

Часопис Инфотека

+381 11 3370-211, infotheca@unilib.rs

11000 Београд, Булевар краља Александра 71

Часопис излази два пута годишње

ШТАМПА:

Дунав

Београд

Садржај

Научни радови

Јована Ковачевић, Јелена Граовац	
Н-грамски заснована класификација текста на српском језику применом методе структуралних подржавајућих вектора	5
Миљана Младеновић	
Онтолошко препознавање реторичких фигура	25
Тања Ивановић	
Лексичка анализа двочланих термилолошких синтагми електродистрибутивног система . . .	49
Милена Милинковић	
Библиометријска и цитатна анализа часописа <i>SPATIUM</i>	73

Стручни радови

Гордана Недељков	
Електронске књиге и нова димензија читања .	99
Милена Обрадовић, Александра Арсенијевић, Михаило Шкорић	
Израда мултимедијалног документа “YU рок сцена”	113

Прикази

Владан Девеџић, Милан Крстић	
Рачунарство у друштвеним наукама: приказ студијског програма мастер студија	130
Јелена Митровић	
Приказ ЕУРОЛАН 2015 летње школе из Рачунарске лингвистике	136

Н-грамски заснована класификација текста на српском језику применом методе структуралних подржавајућих вектора

УДК 811.163.41'322.2

САЖЕТАК: У раду су представљени резултати класификације хијерархијски организованог корпуса докумената на српском језику коришћењем методе подржавајућих вектора (МПВ, енгл. *Support Vector Machine, SVM*). Примењене су две технике класификације изведене из методе МПВ са структурним излазом: вишекласна равна (енгл. *flat*) и хијерархијска класификација. Модел заједничке репрезентације документа и класе или хијерархије класа којима документ припада, специфичан за овај облик МПВ методе, базиран је на н-грамима бајтова различите дужине. Коришћене су четири tf-idf статистике које одређују значајност н-грама за одређени документ. Описане технике и статистике тестиране су на хијерархијски структурираном подскупу Ебарт корпуса новинских текстова. Добијени резултати за оба типа класификатора на нивоу целог корпуса су приближни, док на нивоу појединачних класа хијерархијски тип класификатора показује боље резултате за већину класа са малим бројем текстова.

КЉУЧНЕ РЕЧИ: хијерархијска класификација текста, метода подржавајућих вектора, Ебарт корпус.

РАД ПРИМЉЕН: 22. мај 2015.

РАД ПРИХВАЋЕН: 30. октобар 2015.

Јована Ковачевић
jovana@matf.bg.ac.rs

Јелена Граовац
jgraovac@matf.bg.ac.rs

Универзитет у Београду
Математички факултет
Катедра за рачунарство
и информатику

1. Увод

Класификација текста је један од задатака истраживања текста (енгл. *text mining*) – области рачунарске лингвистике која обухвата скуп техника за издавање корисне, скривене, претходно непознате информације из текстуалних

докумената. У случају класификације текста скривена информација је припадност текста, према садржају, једној или већем броју класа из предефинисаног скупа класа (Manning and Schütze, 1999). Класификација може да се спроведе ручно, али је тај поступак временски захтеван и скуп. С обзиром на широку распрострањеност брзих рачунара, аутоматска класификација постала је основ за ефикасну обраду великих колекција докумената и откривање знања садржаног у њима.

У аутоматској или полуаутоматској класификацији текста углавном се користе два различита приступа: приступ заснован на лексичко-семантичким језичким ресурсима и приступ заснован на машинском учењу. Системи прве врсте користе лексичко-семантичке мреже, као што су WordNet (Miller, 1995) и FrameNet (Johnson et al.), уз ресурсе и алате какви су електронски речници и лексикон граматике (Gross, 1997), семантичке онтологије, онтологије именованих ентитета и властитих имена. Ови језички ресурси омогућају развој модела класификације за морфолошки и деривационо изузетно богат језик какав је српски. Класификатор се најчешће ручно гради на основу богатих морфолошких, синтаксичких и семантичких информација садржаних у лексичким ресурсима (Scott and Matwin, 1998) и не захтева се постојање скупа класификованих текстова који би се користили за учење класификатора. За разлику од ових, системи који користе други приступ подразумевају постојање класификованих корпуса текстова подељених на скуп за учење и скуп за тестирање. На основу података за учење граде се класификатори (класификациони модели, функције припадности класи) применом разних статистичких метода, нпр. бајесовске класификације (Eyheramendy et al., 2003), условних случајних поља (McCallum and Pereira, 2001), скривених Марковљевих модела (Yi and Beheshti, 2013), или метода заснованих на подржавајућим векторима, неуронским мрежама (Sebastiani, 2002), најближим суседима (Yang and Pedersen, 1997), дрветима одлучивања (Quinlan, 1996), итд. Посебно је значајан вишејезични EuroVoc класификатор JEX (Steinberger et al., 2013) који се састоји од обучених класификатора за 22 различита језика Европске уније.

Одабир методе, па и самог приступа проблему класификације текста зависи од два кључна фактора: доступности језичких ресурса и доступности података за обучавање. Уколико језички ресурси, као што су лексикони, речници и граматике, семантичке мреже постоје, има смисла користити приступ заснован на њима. Овај приступ мора да узме у обзир карактеристике и специфичности сваког језика на који се примењује, па тако у случају српског језика, чињеницу да користи два алфабета (ћирилични и латинични), да је правопис фонолошки заснован, да је морфолошки систем богат, ред речи у реченици слободан, систем слагања веома сложен (Vitas et al., 2003). Све ове карактеристике чине да су кораци претходне обраде, као што су избор и издвајање својстава на основу

којих се врши класификација (енгл. *feature selection, feature extraction*) доста сложени. Уколико корпуси потребни за обучавање алгоритама – у случају класификације текстова то су базе класификованих текстова – постоје или их је лако и јефтино направити, погодно је применити приступ заснован на машинском учењу. Техникама машинског учења класификатор се генерише аутоматски, „учењем“ карактеристика класа на основу скупа података за учење придружених свакој класи. То су подаци који су ручно класификовани у класе од стране експерта из домена. Након процеса подучавања (тренирања, учења), класификатор најчешће аутоматски генерише скуп правила која треба да задовољава податак да би био класификован у одређену класу.

У зависности од броја класа, класификација може бити бинарна, када су дефинисане само две могуће класе или вишекласна, када је дефинисано више могућих класа. У зависности од тога да ли се класе могу преклапати или не, класификација може бити једнозначна (енгл. *single-label*), када једном податку може бити додељена тачно једна класа или вишезначна (енгл. *multi-label*), када једном податку може бити додељена једна, ни једна или више класа, односно класе се могу преклапати. Према структури која дефинише односе међу класама, класификација може бити хијерархијска или нехијерархијска. Уколико се у току процеса класификације класе посматрају самостално без икакве структуре која дефинише односе између њих, тада се ради о нехијерархијској класификацији. Када број различитих класа, или број података унутар једне класе постане јако велики, јављају се проблеми тачног и ефикасног претраживања и управљања подацима на нивоу једне класе. У том случају, класе се најчешће организују у стаблике структуре и уводи се хијерархијска структура међу њима (Sun and Lim, 2001) (нпр. Yahoo хијерархија).

У (Graovac, 2013) и (Pavlović-Lazetić and Graovac, 2010) приказана је метода класификације докумената на српском језику заснована на српском WordNet-у (Krstev et al., 2004) развијеном за српски језик у оквиру Групе за језичке технологије на Математичком факултету Универзитета у Београду¹ и примењена је на корпус новинских текстова Ебарт². Коришћени су и други језички ресурси развијени за српски језик у оквиру Групе – електронски речник (Vitas and Krstev, 2005), лексикон граматике (Vitas et al., 2003), онтологија властитих имена (Krstev et al., 2005).

У радовима (Graovac et al., 2015; Graovac, 2014a; Graovac and Pavlović-Lazetić, 2014; Graovac, 2014b) приказане су методе класификације засноване на машинском учењу које користе n -грамску методу за представљање текста и методу k најближих суседа (енгл. *k nearest neighbors, kNN*) за изградњу класификатора. Методе су језички независне, примењене су на корпусе текстова

¹ [www.matf.bg.ac.rs/\\$\sim\\$scvetana/LT-pregled.html](http://www.matf.bg.ac.rs/$\sim$scvetana/LT-pregled.html)

² <http://www.arhiv.rs/novinska-arhiva/>

најраспрострањенијих писама и језика, различитих лексичких, морфолошких, синтаксичких и правописних карактеристика (енглески, кинески, арапски, шпански), веома су једноставне и показале су веома добре резултате. Посебно, у (Граовац, 2012) приказана је примена н-грамске методе машинског учења на класификацију текстова на српском језику и корпусу новинских текстова Ебарт. Класификација је вишекласна, вишезначна и нехијерархијска. У овом раду по први пут је примењена метода структуралних подржавајућих вектора на класификацију текстова на српском језику (Ебарт корпус). Над „равним“ (енгл. *flat*) корпусом дефинисана је хијерархијска структура класа а из самог корпуса издвојен је подскуп докумената који по садржају одговарају тим класама. Затим су над овим хијерархијски организованим корпусом примењене две технике класификације, изведене из методе МПВ са структуралним излазом: вишекласна (равна, *flat*) класификација (избор једне од већег броја класа) и хијерархијска класификација (избор хијерархије класа којима документ припада). Модел заједничке репрезентације документа и класе или хијерархије класа којима документ припада, специфичан за овај облик МПВ методе, базиран је на н-грамима бајтова различите дужине.

Најзад, на основу података за тестирање евалуиран је модел и одређена је тачност резултата коришћењем једне од стандардних мера у области претраживања информација – $F1$ мере која комбинује одзив (енгл. *recall*) и прецизност (енгл. *precision*) (Tan et al., 2006).

У даљем тексту рада биће прво приказан Ебарт корпус – највећа дигитална медијска документација у Србији, као и хијерархијски организован подкорпус издвојен из Ебарт корпуса за потребе тестирања метода хијерархијске класификације (део 2.).

Затим ће бити приказана у основним цртама коришћена методологија (део 3.): метода МПВ са структурним излазом и њена прилагођавања за примену на вишекласну и хијерархијску класификацију (тачка 3.1), концепт н-грама бајтова (тачка 3.2) и специфични н-грамски начин заједничког представљања документа и класе (односно хијерархије класа) којој документ припада као и кораци учења и тестирања у примени ове методе (тачка 3.3). Мере евалуације уведене су у тачки 3.4.

Главни резултат рада – резултат класификације текстова – биће приказан у делу 4.. Биће приказани и резултати евалуације и однос према сродним (упоредивим) резултатима. Најзад, у делу 5. интерпретираћемо и дискутовати добијене резултате, утицај примењених метода и могућности усавршавања.

2. Подаци (Dataset)

Ебарт (www.arhiv.rs/novinska-arhiva/) представља највећи архив новинских текстова савременог српског језика у дигиталном облику. Он постоји од 2003. године и до данас је у њему ускладиштено више од 2.000.000 текстова из штампаних медија. Актуелна архива је класификована на тематске целине по угледу на уобичајене новинске рубрике: политика, спољна политика, друштво, економија, хроника, култура, забава, спорт, медији, фељтон, писма читалаца, и друго. У циљу тестирања методе подржавајућих вектора на проблему хијерархијске класификације текстова, из „равног“ Ебарт корпуса је издвојен хијерархијски организован подкорпус који смо назвали ЕбартХиер. ЕбартХиер обухвата све чланке из дневног листа Политика, објављене од 2003. до 2006. године, који припадају рубрикама Политика, Друштво, Економија и бизнис, Светска привреда и финансије, Култура, Наука и технологија, као и чланке из листа Спортски журнал (објављене од 2003. до 2006) који припадају рубрикама Кошарка и Фудбал. Сви документи који припадају по тематици сличним рубрикама/класама, груписани су у заједничку класу на вишем нивоу хијерархије. Тако су настале класе Политика и друштво, Економија, Култура и наука и Спорт. Овако добијени корпус се карактерише дрволиком структуром приказаном на Слици 1.



Слика 1. Приказ дрволике структуре ЕбартХиер корпуса. Сваком имену класе придружен је број којим се та класа нумерише. У листовима, у угластим заградама, приказан је број докумената у свакој класи

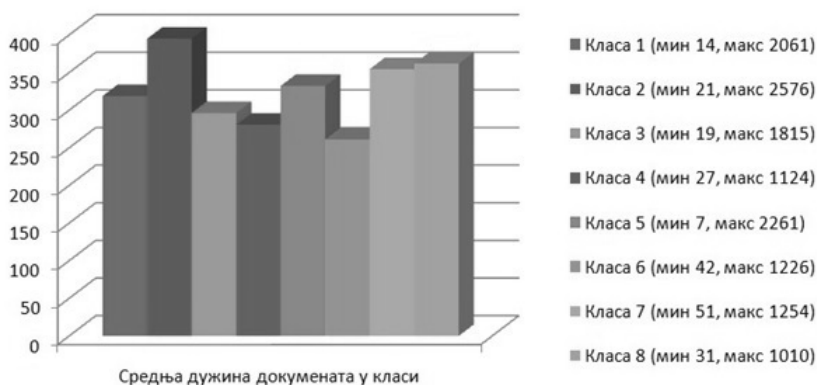
Класе су непреклапајуће (један документ може припадати само једној класи) и сваки документ се може класификовати само у класу која припада листу хијерархије. Корпус се карактерише изразито неравномерном расподелом докумената по класама (видети Сliku 1). Средња вредност дужине свих докумената у корпусу је 327,75 речи у документу (најкраћи документ има 7

а најдужи 2576 речи). На Слици 2 је приказан хистограм средњих вредности дужина докумената (као и дужине најкраћег и најдужег документа) по свим класама. Сви документи су представљени латиничним писмом коришћењем UTF-8 кодне шеме.

3. Методологија

3.1 Метод структуралних подржавајућих вектора

Метода подржавајућих вектора (МПВ) се показала веома ефикасном у класификацији текстова (Joachims, 1998). У овом раду биће примењена метода структуралних подржавајућих вектора (МСПВ, (Tsochantaridis et al., 2004)) која представља генерализацију МПВ методе на структурални излаз, који може бити у облику низа, стабла, усмереног ацикличног графа итд.



Слика 2. Хистограм средњих дужина докумената у класама. У заградама су додатно приказане дужине најкраћег и најдужег документа у свакој класи

Пре него што представимо МСПВ методу, биће дат преглед основне МПВ методе за бинарну класификацију из угла класификације текстова. Стандардни приступ у обучавању предиктора за бинарну класификацију је учење дискриминантне функције (енгл. *Discriminant function*) $F(x)$ и класификовање улазног вектора x на основу знака функције $F(x)$. С обзиром да линеарне методе обично имају ефикасне алгоритме за обучавање, уобичајено је

претпоставити да је функција $F(x)$ линеарна, односно да се може представити у облику $F(x) = \langle \omega, x \rangle$, где је ω вектор параметара који се уче а „ $\langle \dots \rangle$ “ ознака за скаларни производ. Улазни вектор x може помоћу неке функције Ψ бити пресликан у други простор и у том случају дискриминантну функцију записујемо $F(x) = \langle \omega, \Psi(x) \rangle$.

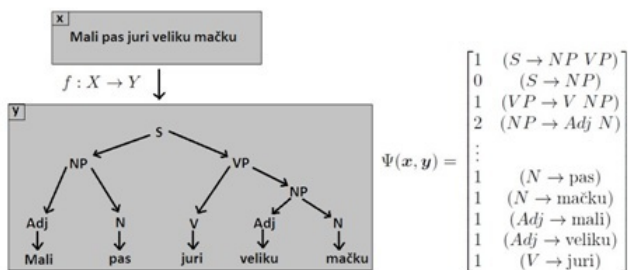
Бинарни класификатор може да предвиди да ли дати текст припада одређеној класи или не, што значи да његов излаз може бити -1 , уколико текст не припада датој класи, или 1 , у супротном. Уколико желимо да знамо којој класи текст припада из датог скупа класа, okreћемо се методама које предвиђају структурални излаз, конкретно МСПВ. У основној МПВ методи закључивање о излазу y за дати улаз x се врши на основу знака дискриминаторне функције, односно $\text{sgn}(F(x)) = y$, где је $y \in \{-1, 1\}$. Било би идеално када бисмо и за структурални случај могли да пронађемо функцију $F(x)$ која тачно пресликава скуп улазних података (у овом случају текстова) у скуп излазних података (у овом случају класа). С обзиром да је веома тешко конструисати баш такву функцију, идеја је конструисати функцију $F : X \times Y \rightarrow \mathbb{R}$ која би мерила колико добро излаз y одговара улазу x .

Желимо да функција F буде таква да што је веће $F(x, y)$, то излаз y боље одговара улазу x , односно у конкретном случају, дата класа боље одговара датом тексту. У овој генерализацији, дискриминантна функција постаје функција два аргумента, улаза и излаза, $F(x, y)$, при чему излаз није из скупа $\{-1, 1\}$ већ може представљати низ, стабло, граф итд. Ако са Y означимо скуп свих могућих излаза, без обзира у ком је облику излаз (низ, стабло, граф, ...), МСПВ предвиђа излаз (класу) који најбоље одговара улазу (тексту) тј. предвиђа излазни вектор у који максимизира вредност функције F за дати улазни вектор x . Прецизније, МСПВ предвиђа излаз на основу следеће једначине: $y^* = \arg\max_{y \in Y} (F(x^*, y))$. Аналогно са МПВ, и у МСПВ претпостављамо

да је функција F линеарна по ω , као и да се векторски пар улаз-излаз (x, y) може погодније представити пресликавањем помоћу неке функције Ψ , у неки други простор, па стога функцију F можемо записати као $F(x, y) = \langle \omega, \Psi(x, y) \rangle$.

Функција Ψ представља заједнички, векторски запис за један пар (x, y) и њен облик зависи од скупа података над којима се метода примењује. На пример, једна од примена МСПВ јесте одређивање („предвиђање“) дрвета извођења дате реченице у датој (формалној) граматичи и у том случају, улазни податак x би представљао вектор речи које се појављују у реченици, излазни податак би представљао дрво извођења у датој граматичи, а функција $\Psi(x, y)$ би била заједничка репрезентација реченице и њеног дрвета извођења. Та заједничка репрезентација би могла да буде вектор чија димензија одговара укупном броју правила у граматичи, укључујући и правила извођења свих речи из скупа за

обучавање. Сваки елемент овог вектора би одговарао једном од свих могућих правила те граматике, а вредност на свакој позицији вектора одговарала би укупном броју појављивања (у дрвету извођења) правила које одговара тој позицији. Пример репрезентације вектора $\Psi(x, y)$ у овом случају приказан је на Слици 3.



Слика 3. Пример заједничке репрезентације улаза и излаза конструисан по узору на сличан пример извођења у енглеској граматичи из (Tsochantaridis et al., 2004)

У примеру приказаном на Слици 3, вектор x чине речи дате реченице („Мали пас јури велику мачку“), а вектор y дрво извођења у датој формалној граматичи. Вектор $\Psi(x, y)$ означава да је у извођењу реченице правило $S \rightarrow NPVP$ примењено једанпут, правило $S \rightarrow NP$ ниједном, правило $VP \rightarrow VNP$ једанпут, правило $NP \rightarrow AdjN$ два пута и тако даље, правило $V \rightarrow juri$ једанпут.

Постоје различите формулације МСПВ методе (Tsochantaridis et al., 2005), а у овом раду коришћена је такозвана формулација са једном „лабавом“ (енгл. *slack*) променљивом са рескалирањем маргине (Joachims et al., 2009). Како је при подучавању класификатора добро да маргина која раздваја инстанце различитих класа буде што шира а како је, са друге стране, могуће да неке инстанце при томе буду погрешно класификоване, ова формулација МСПВ подучава параметре класификатора у зависности од једне позитивне константе C која контролише нагодбу између минимизације грешке над скупом за подучавање и максимизације маргине (ова константа биће посебно анализирана у даљем раду). Описани алгоритам има полиномијалну сложеност по броју примера за подучавање, што је доказано у (Joachims et al., 2009).

3.2 Н-грами

Ако је дата ниска симбола $s = s_1 s_2 \dots s_N$ над азбуком Σ , н-грам ниске s (за N , n природне бројеве) дефинише се као било која подниска суседних симбола ниске s дужине n . Над азбуком Σ може се дефинисати укупно $|\Sigma|^n$ различитих н-грама, при чему је $|\Sigma|$ величина (кардиналност) азбуке Σ . Овако представљени н-грами могу бити дефинисани на нивоу речи, карактера или бајта. На пример, 2-грами (уобичајен назив биграми) над ниском „dela, ne reči“ на нивоу речи садржаће само два биграма, „dela ne“ и „ne reči“ а биграми карактера (азбука Σ је латинична), биће de; el; la; a.; , _; _n; ne; e _; _r; re; eč; či. Ако су карактери кодирани UTF-8 кодном схемом, слово „č“ кодирано је са два бајта чији је декадни садржај 196 141, редом, карактер „ ” (space, белина) кодиран је једним бајтом чији је садржај декадни број 32, итд, па је цела ниска у рачунару записана низом бајтова чија је декадна вредност 100 101 108 97 44 32 110 101 32 114 101 196 141 105. Дакле, у случају језика над латиничном азбуком, н-грами на нивоу бајта и нивоу карактера су веома слични с обзиром на чињеницу да је један карактер обично представљен једним бајтом. Разлика је обично и у скупу карактера који се разматрају (н-грами на нивоу карактера обично не узимају у обзир разлику између малих и великих слова, интерпункцију, цифре и размаке), а разлика је посебно значајна када се користи ћирилично писмо или друга писма као што су арапско, кинеско, и сл. Н-грами бајтова и н-грами карактера се равноправно користе за репрезентацију текстова у решавању различитих задатака из области истраживања података (Kešelj et al., 2003; Abou-Assaleh et al., 2004; Reddy and Pujari, 2006; Santos et al., 2011; Lui et al., 2014), са сличним резултатима. Мада н-грами бајтова понекад немају придружен смисао, посебно за човека (на пример, када садрже само један од два бајта којима је представљен један карактер), за њихову екстракцију није неопходно поседовати информацију о кодној шеми која је коришћена за запис текста па зато они представљају поједностављену репрезентацију за рад рачунара. У овом раду ми ћемо користити н-граме бајтова.

Када се користе у процесу обраде природних језика, неке од добрих особина које н-грами показују су релативна неосетљивост на правописне грешке, азбука знакова је унапред позната, независност од језика и садржаја, извршавање у једном пролазу, не захтева се никакво лингвистичко предзнање, и сл. Основни проблем код коришћења н-грама је експоненцијални број њихових могућности у односу на кардиналност азбуке. Ако је Σ азбука енглеског језика и ако се у разматрање укључи и знак за празнину, онда је $|\Sigma| = 27$. Ако се прави разлика између малих и великих слова и ако се у разматрање укључе цифре, онда је $|\Sigma| = 63$. Јасно је да ће многи алгоритми са н-грамима бити веома скупи са становишта израчунљивости већ за $n = 5$ или $n = 6$ (на пример, број различитих 5-грама над азбуком Σ је $63^5 \sim 10^9$). Коришћење модела и

техника заснованих на н-грамима у процесу обраде природних језика показало се као ефикасан приступ. Овај приступ нашао је примену у оквиру задатка претраживања информација (De Neer, 1974), компресије текста (Wiśniewski, 1987), откривања и исправљања правописних грешака (Zamora et al., 1981), откривање ауторства текста (Kešelj et al., 2003) и друго.

3.3 Репрезентација података

За потребе овог истраживања развијена су два класификатора који за основу имају оригинални МСПВ алгоритам и међусобно се разликују по начину на који су излазни подаци, односно класе текстова, представљени. Оба класификатора користе исту репрезентацију улазних података, н-грамима на нивоу бајта. Свака позиција у вектору н-грама једнака је вредности *tf-idf* статистике за дати н-грам. Статистика *tf-idf* (енгл. *term frequency – inverse document frequency*) је обично дефинисана тако да рефлектује колико је неки н-грам важан за документ у оквиру неког корпуса. Ова мера пропорционално расте са порастом броја појава н-грама у документу, али опада са порастом броја његових појава у целом корпусу. За документ ће бити значајнији они н-грами који имају већу вредност *tf-idf* статистике, односно они н-грами који се чешће јављају у том документу али се не јављају често у целом корпусу. Постоје разне варијанте за рачунање вредности *tf* и *idf*, а у овом раду су коришћене следеће мере (Manning et al., 2008)

1. *classic tf-idf*: $tf \cdot \log\left(\frac{n}{n_k+1}\right)$
2. *log tf-idf*: $1 + \log(tf) \cdot \log\left(\frac{n}{n_k+1}\right)$
3. *boolean1 tf-idf*: $\log\left(\frac{n}{n_k+1}\right)$
4. *boolean2 tf-idf*: $\log\left(1 + \frac{n}{n_k}\right)$

при чему *tf* представља нормализовану фреквенцију н-грама у припадајућем документу, *n* представља укупан број докумената у целом корпусу а *n_k* представља број оних докумената у корпусу у којима се бар једном појављује посматрани н-грам.

Репрезентација излазних података односно класа код развијених класификатора се разликује. Сваком тексту ЕбартХиер корпуса придружена је једна класа која се налази у листу ЕбартХиер хијерархије. У првом класификатору, свака класа је представљена као јединствени природни број, не узимајући притом у обзир повезаност класа кроз хијерархију. Ако класе нумеришемо редом као на Слици 1, скуп *Y* сводимо на скуп {1, 2, 3, 4, 5, 6, 7, 8}. На овај начин основну МСПВ методу сводимо на вишекласни равни класификатор. Вектор $\Psi(x, y)$, заједничка репрезентација улазног и излазног податка, у првом класификатору је димензије $p \cdot q$, где је *p* укупан број

различитих n -грама, односно димензија вектора x , а q купан број различитих класа корпуса. На тај начин, свака класа добија свој блок у ком се налазе нуле, уколико дати текст не припада датој класи, или вредности улазног вектора x , у супротном. На пример, ако текст x припада класи k , тада ће заједничка репрезентација имати следећи облик:

$$\Psi(x, y) = \left[\underbrace{0, \dots, 0}_{\text{класа 1}}, \dots, \underbrace{x_1, \dots, x_p}_{\text{класа k}}, \dots, \underbrace{0, \dots, 0}_{\text{класа q}} \right]$$

У другом класификатору, свака класа је представљена као вектор чворова у ЕбартХиер хијерархији. Сваком чвору одговара једна позиција у вектору, при чему се на некој позицији налази „1“ ако се дати чвор појављује у путањи од корена хијерархије до листа дате класе, а „0“ у супротном. У нашем примеру (Слика 1), класа економија и бизнис била би представљена као $(0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1)$. На овај начин МСПВ методу сводимо на хијерархијски класификатор.

Вектор $\Psi(x, y)$ је у другом класификатору димензије $p \cdot r$, где је p укупан број различитих n -грама, односно одговара димензији вектора x , а r укупан број различитих чворова хијерархије. На тај начин, сваки чвор добија свој блок у коме се налазе нуле, уколико дати текст не припада класи која садржи дати чвор, или вредности улазног вектора x , у супротном. На пример, ако текст x припада класи економија и бизнис, представљеној са $(0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1)$, тада ће заједничка репрезентација имати следећи облик $\Psi(x, y) = [0, 0, x, 0, 0, 0, 0, 0, 0, x, 0, 0, x]$ где је $0 = \underbrace{[0, \dots, 0]}_{r \text{ пута}}$, а вектор

x одговара улазном вектору. Основна разлика у репрезентацији вектора $\Psi(x, y)$ између равног и хијерархијског класификатора је у томе што равни класификатор третира све класе појединачно, док хијерархијски узима у обзир да су класе део дрволике хијерархије. Ова разлика се осликава на заједничку репрезентацију улазног и излазног податка, вектор $\Psi(x, y)$, који је различитих димензија за два случаја: код равног класификатора резервисане су позиције за сваку класу (која се налази у листовима хијерархије) док су код хијерархијског класификатора резервисане позиције за сваки чвор хијерархије класа, укључујући и листове, и унутрашње чворове и корен.

3.4 Опис имплементације

Оба класификатора су настала прилагођавањем јавно доступног и бесплатног оквира за МСПВ методу SVM^{struct}³. SVM^{struct} је доступан у разним

³ http://www.cs.cornell.edu/people/tj/svm_light/svm_struct.html

програмским језицима а овде је била коришћена имплементација у програмском језику С. Прилагођавање основне имплементације подразумевало је, између осталог, допуњавање постојећих структура за улазни вектор x и излазни вектор y , имплементацију функције за генерисање вектора $\Psi(x, y)$ (заједничке репрезентације улазног вектора x и излазног вектора y), имплементацију функције која дефинише меру тачности и имплементацију функције за евалуацију квалитета класификатора.

Подаци су припремани на следећи начин:

1. за сваки текст је алатом Text::Ngrams (Kešelj et al., 2003) генерисан низ н-грама бајтова које он садржи као и колико се пута сваки н-грам бајтова појављује. Генерисани су подаци за н-граме бајтова дужине $\{2, 3, 4, 5, 6, 7\}$.
2. имплементиран је скрипт који је за сваки текст на основу н-грама генерисао улазни вектор x у формату који захтева класификатор SVM^{struct}.

3.5 Опис експеримента

Целокупан корпус текстова подељен је на скуп за подучавање и скуп за тестирање у односу 2:1, што је један од најчешћих начина поделе у класификацији текстова (Bellotti and Crook, 2009). Од укупно 60.637 текстова, скуп за подучавање је сачињавао 40.426 а скуп за тестирање 20.211 текстова. Расподела текстова по класама такође прати ову размену и приказана је у Табели 1.

Класа	Скуп за подучавање	Скуп за тестирање	Цео корпус
1	10681	5340	16021
2	7173	3586	10759
3	12290	6145	18435
4	772	386	1158
5	9172	4586	13758
6	140	70	210
7	32	15	47
8	166	83	249
Укупно	40426	20211	60637

Табела 1. Број текстова у скупу за подучавање и скупу за тестирање по класама

Сваки текст ЕбартХиер корпуса представљен је н-грамима бајтова на основу којих се изграђује векторска репрезентација погодна за улаз у класификатор,

такозвани улазни вектор x . Сваком n -граму бајтова из корпуса одговара једна позиција у вектору x , а вредност на тој позицији једнака је вредности једне од 4 tf-idf статистике за одговарајући n -грам бајтова. У зависности од дужине n -грама бајтова ($n \in \{2 \dots 7\}$) и од одабране tf-idf мере (classic, log, boolean1, boolean2), генерисано је 24 репрезентације корпуса ($n = 2$, мера=classic, ..., $n = 7$, мера=classic, $n = 2$, мера=log, ..., $n = 7$, мера=log, ..., $n = 2$, мера=boolean1, ..., $n = 7$, мера=boolean1, $n = 2$, мера=boolean2, ..., $n = 7$, мера=boolean2). У свакој репрезентацији, скупови за подучавање и тестирање се састоје од истих текстова како би се могли међусобно поредити.

У намери да испитамо да ли један тип класификатора даје боље резултате за одређену дужину n -грама бајтова или за одређену tf-idf статистику, извршили смо равну и хијерархијску класификацију за сваку репрезентацију корпуса на следећи начин:

1. Извршена је 10-унакрсна валидација (енгл. *cross validation*) на скупу за подучавање којом је одређена оптимална вредност параметра C .
2. За добијену вредност параметра, подучаван је класификациони модел на целом скупу за подучавање.
3. Тестирање и евалуација модела спроведени су на скупу за тестирање.

Након добијених резултата тестирања, извршено је поређење перформанси добијених модела.

3.6 Евалуација

Да би се анализирале перформансе МСПВ метода хијерархијске класификације, користимо неке од уобичајених мера квалитета класификације – прецизност (енгл. *precision*), која се дефинише као проценат примера који су исправно класификовани међу свим примерима који су додељени одређеној класи, одзив (покривање, енгл. *recall*) који дефинише колико примера за тестирање из дате класе класификатор може да препозна и F -мера (енгл. *F-measure*, $F1$), која се дефинише као хармонијска средина прецизности и одзива (Baeza-Yates et al., 1999):

$$Precision = \frac{TP}{TP + FP} \quad Recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

при чему је TP (True Positives) број тачно позитивно класификованих докумената, TN (True Negative) број тачно негативно класификованих докумената, FP (False Positive) број погрешно позитивно класификованих

докумената и FN (False Negative) број погрешно негативно класификованих докумената.

Ове мере су дефинисане за случај бинарне класификације (када постоје само две класе). У случају када се класификација врши на више од две класе, потребно је извршити усредњавање ових мера по свим класама. То може да буде урађено на два начина: може се тражити макроресек, где се свакој класи придаје исти значај, и микроресек, где се фаворизују класе које садрже већи број докумената. Код макроресека се најпре израчунава вредност мере за сваку класу појединачно, а затим се врши усредњавање тих вредности по броју класа. Код микроесека се израчунавају вредности за TP , TN , FP и FN за сваку класу појединачно. Затим се израчунавају вредности TP , TN , FP и FN као суме свих TP , TN , FP и FN за све класе редом. На крају се израчунава вредност мере за добијене сумиране вредности TP , TN , FP и FN . У овом раду користимо микроресек Φ -мере (микро- Φ 1). Основни недостатак ових мера је подједнако кажњавање грешака на разним нивоима хијерархије.

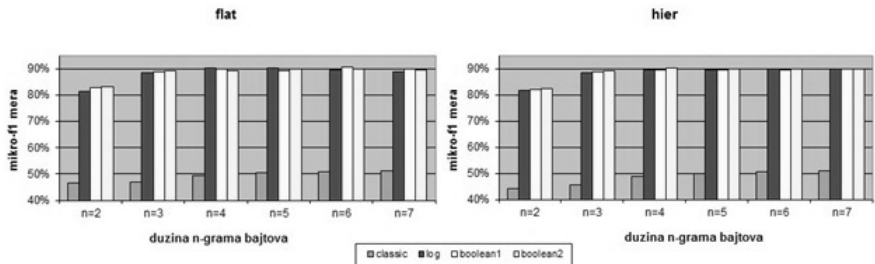
4. Резултати

На ЕбартХиер корпус примењене су две варијанте МСПВ методе: у једној се спроводи равна класификација, тј. не узима се у обзир да су класе хијерархијски повезане, а у другој се спроводи хијерархијска класификација. За текстове корпуса генерисано је 6 различитих н-грамских репрезентација, за н-граме бајтова дужине n из скупа $\{2, \dots, 7\}$, за 4 различите $tf-idf$ мере приказане у поглављу 3.3. Над сваким овако добијеним скупом за подучавање, за оба типа класификатора, извршена је 10-унакрсна валидација којом је одређен параметар класификатора C из скупа вредности 10^{-2} до 10^2 са кораком 10, који даје најбоље резултате.

На Слици 4 приказани су резултати евалуације оба типа класификатора, представљени микро- Φ 1 мером, за корпус представљен н-грамима бајтова дужине од 2 до 7 и за све 4 $tf-idf$ мере, док су комплетни нумерички подаци приказани у Табели 2. Оба типа класификатора (равни и хијерархијски) показују тенденцију да класификатори са мером *classic* имају лошије перформансе од осталих. Вредности микро- Φ 1 мере су за остале мере приближне, при чему разлике микро- Φ 1 мере различитих класификатора за исту меру скоро нигде не прелазе 1% (осим за меру *classic*, за $n=2$ и $n=3$ где редом износе 2,32% и 1,66%).

Перформансе најбољих класификатора за оба типа (равни и хијерархијски) анализиране су и по класама. Међу равним класификаторима, највећу вредност микро- Φ 1 мере имао је класификатор за меру *boolean1* за текстове представљене 6-грамима бајтова (микро- Φ 1мера износи 90,43%) а за хијерархијски за меру

boolean2 за текстове представљене 4-грамима бајтова (микро-Ф1 мера износи 90,17%). С обзиром да је број текстова по класама различит (наведен на Слици 1), класе се могу поделити у две групе: „велике“ (1, 2, 3 и 5) и „мале“ (4, 6, 7, 8). Резултати су представљени на Слици 5. Резултати за „велике“ класе су приближни за оба типа класификатора, док су разлике знатно израженије за мале класе. Најбољи хијерархијски класификатор много боље предвиђа „мале“ класе 6 и 7. Највећу тачност најбољи класификатори оба типа показују за „малу“ класу означену са 8 (фудбал).



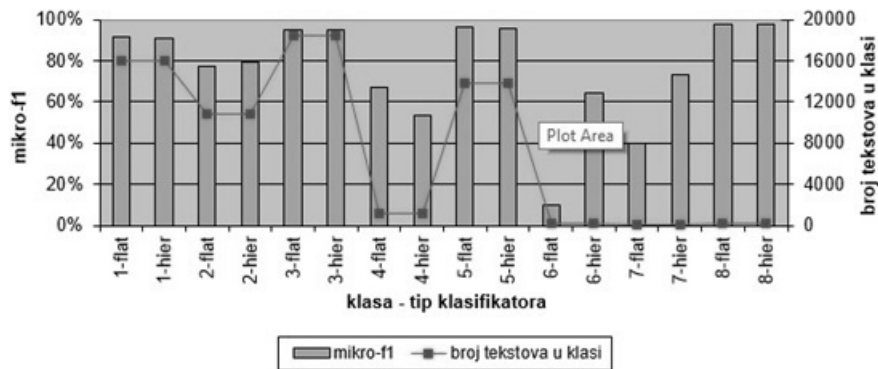
Слика 4. Резултати евалуације два типа класификатора (flat-равног и hier-хијерархијског) за различите улазне податке (дужине n-грама бајтова од 2 до 7) и различите мере

flat	classic	log	boolean1	boolean2	hier	classic	log	boolean1	boolean2
n=2	46.53%	81.46%	82.71%	83.25%	n=2	44.21%	81.75%	82.04%	82.45%
n=3	47.08%	88.27%	88.70%	89.28%	n=3	45.42%	88.34%	88.86%	89.24%
n=4	49.33%	90.05%	90.02%	89.18%	n=4	48.61%	89.39%	89.58%	90.17%
n=5	50.35%	90.05%	89.23%	89.77%	n=5	49.84%	89.46%	89.57%	89.70%
n=6	50.89%	89.60%	90.43%	89.74%	n=6	50.65%	89.84%	89.47%	89.78%
n=7	51.12%	88.88%	89.88%	89.60%	n=7	51.01%	89.88%	89.68%	89.69%

Табела 2. Перформансе равног (flat, табела са леве стране) и хијерархијског (hier, табела са десне стране) класификатора за различите репрезентације улазног корпуса приказани мером микро-Ф1

5. Закључак

Класификација текстова на српском језику из хијерархијски организованог корпуса, применом методе подржавајућих вектора показује приближне резултате за равну и хијерархијску варијанту, за сваку tf-idf меру.



Слика 5. Резултати евалуације најбољих класификатора оба типа (flat-равног и hier-хијерархијског) по класама. Приказан је и број текстова у корпусу за сваку класу

Укупно најбољи резултат добијен је равном варијантом класификатора за boolean1 меру и износи 90,43%. Ово је нешто бољи резултат од раније публикованог резултата за н-грамску класификацију равно организованог подскупа Ебарт корпуса за који је израчуната вредност микро-Ф1 мере 88,5% (Граовац, 2012). Хијерархијска класификација даје нешто боље резултате од равне за неке „мале“ класе.

Хијерархијска класификација даје резултате слабије од очекиваних. Разлог за то се може тражити у плиткој хијерархији са малим бројем докумената, али такође и у коришћеној мери евалуације. При евалуацији резултата класификације, посебно хијерархијске, није довољно пребројати промашаје већ је потребно проценити и њихову тежину, односно растојање предвиђене од тачне класе.

Имајући то у виду, очекујемо да ће се бољом, хијерархијској класификацији прилагођеном мером евалуације као и богатијом хијерархијом докумената постићи резултати који превазилазе резултате равне класификације (Silla et al., 2011).

Литература

- Abou-Assaleh, Tony, Nick Cercone, Vlado Kešelj, and Ray Sweidan. "N-gram-based detection of new malicious code". In *Computer Software and Applications Conference, COMPSAC 2004. Proceedings of the 28th Annual International, IEEE*. Vol. 2 (2004), 41-42. Accessed September 1, 2015. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1342667>.
- Baeza-Yates, Ricardo, Ribeiro-Neto Berthier, et al. *Modern information retrieval*, vol. 463. ACM press New York, 1999.
- Bellotti, Tony and Jonathan Crook. "Support vector machines for creditscoring and discovery of significant features". *Expert Systems with Applications*, vol. 36, no. 2 (2009): 3302-3308. Accessed September 1, 2015. <http://www.sciencedirect.com/science/article/pii/S0957417408000857>.
- De Heer, T. "Experiments with syntactic traces in information retrieval". *Information Storage and Retrieval*, vol. 10, no. 3 (1974): 133-144. Accessed September 1, 2015. <http://www.sciencedirect.com/science/article/pii/0020027174900151>.
- Eyheramendy, Susana, David D. Lewis, and David Madigan. "On the Naive Bayes model for text categorization". In *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics Conference*, 332-339. Florida: Society for Artificial Intelligence and Statistics, 2003.
- Graovac, Jelena. "Serbian text categorization using byte level n-grams". In *Proceedings of CLoBL 2012: Workshop on Computational Linguistics and Natural Language, 5th Balkan Conference in Informatics*, 93-97. 2012 Accessed September 1, 2015. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.416.6155&rep=rep1&type=pdf>.
- Graovac, Jelena. "Wordnet-based text categorization technique". *INFOtheca – Journal of Information and Library Science*, vol. 14, no. 2 (2013): 2-17. Accessed September 1, 2015. <http://infoteka.unilib.rs/2013/br.2/eng/Infotheca-2-2013-Jelena-Graovac.pdf>.
- Graovac, Jelena. "A variant of n-gram based language-independent text categorization". *Intelligent Data Analysis*, vol. 18, no. 4 (2014): 677-695.
- Graovac, Jelena. "Text categorization using n-gram based language independent technique". Natural Language Processing for Serbian - Resources and Applications. In *Proceedings of the Conference "35th Anniversary of Computational Linguistics in Serbia"*, 124-135, 2014. Accessed September 1, 2015. <http://poincare.matf.bg.ac.rs/~jgraovac/publications/jg35CLS.pdf>.
- Graovac, Jelena and Gordana Pavlović-Lazetić. "Language-Independent Sentiment Polarity Detection in Movie Reviews: A Case Study of English and Spanish". In *6th International Conference ICT Innovations 2014*, 13-22, 2014. Accessed Sep-

- tember 1, 2015. <http://poincare.matf.bg.ac.rs/~jgraovac/publications/jgSPD.pdf>.
- Graovac, Jelena and Jelena Kovačević, and Gordana Pavlović-Lažetić. “Language Independent n-Gram-Based Text Categorization with Weighting Factors: A Case Study”. *JIDM - Journal of Information and Data Management*, vol. 6, no. 1 (2015): 4–17.
- Gross, Maurice. “The construction of local grammars”. In *Finite State Language Processing eds. Emmanuel Roche and Yves Schabbs*, 329–354. Massachusetts: The MIT Press, 1997. Accessed September 1, 2015. <https://halshs.archives-ouvertes.fr/halshs-00278316/document>.
- Joachims, Thorsten. “Text categorization with support vector machines: Learning with many relevant features”. Springer Berlin Heidelberg, 1998.
- Joachims, Thorsten, Thomas Finely, and Chun-Nam John Yu. “Cutting Plane Training of Structural SVMs”. *Machine Learning Journal*, vol. 77, no. 1 (2009): 27–59. Accessed September 1, 2015. <http://link.springer.com/article/10.1007/s10994-009-5108-8>.
- Johnson, Christopher R., Miriam R. L. Petruck, Collin F. Backer, Michael Ellsworth, Josef Ruppenhofer, and Charles J. Fillmore. “FrameNet: Theory and Practice”, 2002. Accessed September 1, 2015. <http://www.icsi.berkeley.edu/framenet>.
- Kešelj, Vlado, Fuchung Peng, Nick Cercone, and Calvin Thomas. “N-gram-based author profiles for authorship attribution”. In *Proceedings of the conference pacific association for computational linguistics, PACLING*, vol. 3, 255–264. 2003. Accessed September 1, 2015. <http://web.cs.dal.ca/~vlado/papers/pacling03.pdf>.
- Krstev, Cvetana, and Gordana Pavlović-Lažetić, and Ivan Obradović. “Using textual and lexical resources in developing Serbian wordnet”. *Romanian Journal of Information Science and Technology*, vol. 7, no. 1–2 (2004): 147–161. Accessed September 1, 2015. http://xn--c1azn.xn--90a3ac/LicnePrezentacije/ivan_obradovic/Radovi/RJIS_2004.pdf.
- Krstev, Cvetana, Duško Vitas, Denis Maurel, and Mickaël Tran. “Multilingual ontology of proper names”. In *Proceedings of 2nd Language & Technology Conference*, April 21–23, 2005, Poznań, Poland, ed. Zygmunt Vetulani, 116–119. Poznań, Wydawnictwo Poznańskie, 2005. Accessed September 1, 2015. <https://hal.archives-ouvertes.fr/hal-01108242/document>.
- Lui, Marco, Jey Han Lau, and Timothy Baldwin. “Automatic detection and language identification of multilingual documents”. *Transactions of the Association for Computational Linguistics 2* (2014): 27–40. Accessed September 1, 2015. <https://tacl2013.cs.columbia.edu/ojs/index.php/tacl/article/view/86/30>.
- Manning, Christopher, and Hinrich Schütze. *Foundations of Statistical Natural Language Processing*. Cambridge: MIT Press, 1999.

- Manning, Christopher, Prabhakar Raghavan, and Hinrich Schütze. "Scoring, term weighting, and the vector space model". In *Introduction to Information Retrieval*. Cambridge: Cambridge University Press, 2008.
- Lafferty, John, Andrew McCallum and Fernando C. N. Pereira. "Conditional random fields: Probabilistic models for segmenting and labeling sequence data". *Machine Learning*, 282–289. 2001
- Miller, George A. "WordNet: a lexical database for English". *Communications of the ACM*, vol. 38, no. 11 (1995): 39–41. Accessed September 1, 2015. <http://nlp.cs.swarthmore.edu/~richardw/papers/miller1995-wordnet.pdf>.
- Pavlović-Lazetić, Gordana and Jelena Graovac. "Ontology-driven conceptual document classification". In *Proceedings of the International Conference on Knowledge Discovery and Information Retrieval*, 383–386. 2010. Accessed September 1, 2015. <http://poincare.matf.bg.ac.rs/~jgraovac/publications/odcdc.pdf>.
- Quinlan, J. R. "Improved use of continuous attributes in c4. 5". *Journal of Artificial Intelligence Research*, vol. 4 (1996): 77–90. Accessed September 1, 2015. <http://www.jair.org/media/279/live-279-1538-jair.pdf>.
- Reddy, D. Krishna Sandeep, and Arun K. Pujari. "N-gram analysis for computer virus detection". *Journal in Computer Virology*, vol. 2, num. 3 (2006): 231–239. Accessed September 1, 2015. <http://link.springer.com/article/10.1007/s11416-006-0027-8>.
- Santos, Igor, Javier Nieves, and Pablo G. Bringas. "Semi-supervised learning for unknown malware detection". In *International Symposium on Distributed Computing and Artificial Intelligence* vol. 91, 415–422. Berlin Heidelberg, 2011.
- Scott, Sam, and Stan Matwin. "Text classification using wordnet hypernyms". In *Usage of WordNet in Natural Language Processing Systems: Proceedings of the Workshop*, 38–44, 1998. Accessed September 1, 2015. http://www.aclweb.org/website/old_anthology/W/W98/W98-0706.pdf.
- Sebastiani, Fabrizio. "Machine learning in automated text categorization". In *ACM computing surveys (CSUR)*, vol. 34, no. 1 (2002): 1–47.
- Silla J., Carlos N. and Alex A. Freitas. "A survey of hierarchical classification across different application domains". *Data Mining and Knowledge Discovery*, vol.22, no. 1–2 (2011): 31–72. Accessed September 1, 2015. <http://link.springer.com/article/10.1007/s10618-010-0175-9>.
- Steinberger, Ralf, Mohamed Ebrahim, and Marco Turchi. "JRC EuroVoc Indexer JEX-A freely available multi-label categorisation tool". *arXiv preprint arXiv*, vol. 1309, no. 5223 (2013). Accessed September 1, 2015. <http://arxiv.org/ftp/arxiv/papers/1309/1309.5223.pdf>.
- Sun, Aixin, and Ee-Peng Lim. "Hierarchical text classification and evaluation". In *Proceedings 2001 IEEE International Conference on Data Mining*, 521–528. Computer Society, 2001.

- Tan, Pang-Ning, Michael Steinbach, and Vipin Kumar. *Introduction to data mining*, vol. 1. Pearson Education India, 2006.
- Tsochantaridis, Ioannis, Thomas Hofman, Thorsten Joachims, and Yasemin Altun. “Support Vector Machine Learning for Interdependant and Structured Output Spaces”. In *Proceedings of the twenty-first international conference on Machine learning*. 104. New York: ACM, 2004.
- Tsochantaridis, Ioannis, Thomas Hofman, Thorsten Joachims, and Yasemin Altun. “Large Margin Methods for Structured and Interdependant Output Variables”. *Journal of Machine Learning Research*, vol 6, (2005): 1453–1484. Accessed September 1, 2015. http://machinelearning.wustl.edu/mlpapers/paper_files/TsochantaridisJHA05.pdf.
- Vitas, Duško and Cvetana Krstev. “Derivational morphology in an e-dictionary of Serbian”. In *Proceedings of 2nd Language & Technology Conference*. April 21–23, 2005, Poznań, Poland, ed. Zygmunt Vetulani, 139–143. Poznań: Wydawnictwo Poznańskie, Accessed September 1, 2015. http://poincare.matf.bg.ac.rs/~cvetana/biblio/ltc_134_vitas_2.pdf.
- Vitas, Duško, Cvetana Krstev, Ivan Obradović, Ljubomir Popović, and Gordana Pavlović-Lažetić. “An overview of resources and basic tools for processing of Serbian written texts”. In *Proceedings of the Workshop on Balkan Language Resources, 1st Balkan Conference in Informatics, Thessaloniki, Greece*, 97–104, (2003). Accessed September 1, 2015. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.73.8096&rep=rep1&type=pdf>.
- Wiśniewski, Janusz L. “Effective text compression with simultaneous digram and trigram encoding”. *Journal of Information Science*, Vol. 13, No.3, (1987): 159–164.
- Yang, Yiming, and Jan O. Pedersen. “A comparative study on feature selection in text categorization”. In *Proceedings of the 14th International Conference on Machine Learning*, 412–420, 1997. Accessed September 1, 2015. <http://www.surdeanu.info/mihai/teaching/ista555-spring15/readings/yang97comparative.pdf>.
- Yi, Kwan, and Jamshid Beheshti. “A text categorization model based on Hidden Markov models”. In *Proceedings of the 31st Annual Conference of the Canadian Association for Information Science*, 275–287, 2013. Accessed September 1, 2015. <http://www.cais-acsi.ca/ojs/index.php/cais/article/view/420/585>.
- Zamora, E. M., Joseph J. Pollock, and Antonio Zamora. “The use of trigram analysis for spelling error detection”. *Information Processing & Management*, vol. 17, no. 6 (1981): 305–316. Accessed September 1, 2015. <http://www.sciencedirect.com/science/article/pii/0306457381900443>.

Онтолошко препознавање реторичких фигура

УДК 811.163.41'322.2

САЖЕТАК: Аутоматско препознавање реторичких фигура (компарације, ироније, сарказма, хумора, метафоре и сл.) све чешће се користи у задацима обраде природног језика, пре свега за унапређење система класификације текста према осећањима, машинског превођења, али и система који анализирају језичке структуре на различитим нивоима. У овом раду предложена је метода аутоматског препознавања и класификације реторичких фигура из групе тропи која користи правила онтолошког закључивања у онтологији Српски ворднет (SWN). Евалуација методе изведена је над реторичком фигуром *компарација*, а статистичка оцена бинарног класификатора ROC кривом ($AUC = 0,696$) указује да се он може успешно користити у решавању ове врсте задатака. За даље обучавање онтологије SWN, предложена је полу-аутоматска метода учења онтологије повећањем броја и врста релација које могу помоћи у откривању фигуративног говора у текстовима на српском језику.

КЉУЧНЕ РЕЧИ: реторичке фигуре, компарација, класификација помоћу онтологије, WordNet

РАД ПРИМЉЕН: 25. фебруар 2016.

РАД ПРИХВАЋЕН: 9. мај 2016.

Миљана Младеновић

ml.miljana@gmail.com

Универзитет у Београду,

Математички факултет

1. Увод

Аутоматско препознавање реторичких фигура (фигуративног говора) у тексту и њихова анотација није ново поље интересовања у обради природног језика. Једно од првих истраживања у области аутоматског препознавања и интерпретације реторичких фигура које се бави идентификацијом *анафоре* (anaphora resolution) дато је 1964. у раду „A question-answering system for high school algebra a word problems“ (Bobrow, 1964), док је 1991. у раду „met*: A method for discriminating metonymy and metaphor by computer“ (Fass, 1991) предмет аутоматског препознавања била *метафора* и *метонимија*. У литератури су

описане методе аутоматског аотирања *метафоре* (Mason, 2004; Hardie et al., 2007; Koller et al., 2008; Shutova et al., 2013), *анафоре* (Mitkov, 2002; Poesio and Artstein, 2010), *метонимије* (Farkas et al., 2007; Leveling, 2007; Nicolae et al., 2007), *ироније* (Carvalho et al., 2009; Veale, 2012), *сарказма* (Tsur et al., 2010; González-Ibáñez et al., 2011) и др. Међутим, методе анализе осећања у тексту (енг. sentiment analysis — SA) највећим делом разматрају текстове који садрже дословно значење или занемарују утицај фигуративног говора на поларитет осећања у тексту.

У раду (Hao and Veale, 2010) анализа употребе фигуративног говора показала је да одређене реторичке фигуре у тексту делују попут модификатора значења чија је улога у промени поларитета осећања речи или фраза приказана у раду „Sentiment Classification of Movie Reviews Using Contextual Valence Shifters“ (Kennedy and Inkpen, 2006). У општем случају, *модификатори* су речи и изрази који смањују, повећавају или у потпуности мењају поларитет осећања речи или фраза уз које стоје (нпр. поларитет речи *срећан* се може у потпуности променити речју *нимало*). На сличан начин делују и одређене групе реторичких фигура чијом се применом може утицати на поларитет осећања речи или фраза у тексту. По својој дефиницији, фигуре *иронија* и *сарказам* мењају поларитет, *дисфемизам* и *хипербола* појачавају постојећи ниво сентименталне изражајности, док га *литота* и *еуфемизам* умањују. *Метафора*, *метонимија*, *оксиморон* и *компарација* имају сложеније механизме деловања у оба правца промене и јачине и поларитета осећања.

Стога, поступак аутоматске или полуаутоматске идентификације реторичких фигура у тексту може побољшати процес анализе осећања, али и машинског превођења и других апликација заснованих на обради природних језика. На пример, традиционални систем класификације текста на основу осећања (Pang et al., 2002) би на исти начин интерпретирао реченице „Он је брз као зец“ и „Он је брз као пуж“, док би се укључењем процеса препознавања фигуративног говора у другој реченици уочила реторичка фигура иронија, чиме би њен поларитет осећања био у потпуности промењен у односу на дословни. Истраживања објављена у (Reyes and Rosso, 2012b) показала су да се тачност класификације у задатку анализе осећања у тексту може знатно побољшати (са 54% до максималних 89,05%) када се укључе предиктори којима се препознаје фигуративни говор, у односу на скуп предиктора који текст третирају у дословном значењу. Метода машинског учења којом се решава проблем SA унапређена је интегисањем са методом заснованом на правилима којом се препознаје фигуративна употреба језика (Rentoumi et al., 2010). У раду (Williams et al., 2015) разматрана је улога идиома у анализи осећања у тексту. У два експеримента аутори су показали побољшање F-мере класификатора са 45% на 64% и са 46% на 61% када

се као предиктори класификације користе и идиоми који носе фигуративно значење. У раду (Carvalho et al., 2011) резултати анализе осећања изражених у коментарима објављеним на веб порталима новинских вести показују да би 11% коментара посматраног скупа било некоректно означено позитивним поларитетом, уколико би изостала анализа и препознавање појаве реторичке фигуре иронија. Због тога је анализа осећања у текстовима у којима је изражен фигуративни говор нови изазов на пољу обраде природног језика. У 2015. је први пут, на глобалном плану, постављен један семантички задатак ове врсте — задатак семантичке анализе фигуративног говора на Твитеру (International Workshop on Semantic Evaluation — SemEval-2015).¹

Истраживања у области обраде фигуративног говора крећу се у два правца:

- ка побољшању метода класификације поларитета осећања (Reyes and Rosso, 2012b; Rentoumi et al., 2010);
- ка бољем разумевању језичких структура — у раду (Veale and Hao, 2009) упоредне анализе ироничних компарација на енглеском и кинеском језику показале су да је *иронија* лингвистички и културолошки феномен, јер се из посматраног скупа компарација које се примењују у енглеском свега 3–4% могу применити и у кинеском језику.

У истраживању фигуративног говора важну улогу имају семантичке мреже као што је ворднет (енг. WordNet) (Mason, 2004; Barbieri et al., 2015), онтологије (Harris and Di Marco, 2009; Kelly et al., 2010) и лексички ресурси као што су корпуси (Mason, 2004; Hao and Veale, 2010; Reyes and Rosso, 2012a), специјализовани речници жаргона, емотикона и интерпункцијских знакова (Carvalho et al., 2009; González-Ibáñez et al., 2011; Barbieri et al., 2015) и лексикони попут SentiWordNet-a (Rentoumi et al., 2010; González-Ibáñez et al., 2011; Barbieri et al., 2015).

У овом раду најпре је предложена метода аутоматског препознавања реторичких фигура из групе троши која се ослања на онтолошко закључивање, а затим и метода учења онтологије у смислу повећања броја и врста релација у SWN које могу помоћи у откривању таквих фигура у текстовима на српском језику.

2. Методе аутоматског препознавања реторичких фигура

Аутоматско препознавање реторичких фигура у тексту зависи од природе фигуре и језичке структуре коју фигура користи. Да би могао бити изграђен ефикасан систем аутоматског препознавања фигуративног говора, потребно је,

¹ <http://alt.qcri.org/semeval2015/index.php?id=tasks> (Retrieved on March, 23th 2016.)

најпре, формално дефинисати и описати реторичке фигури. Поступак изградње прве формалне доменске онтологије реторичких фигура у српском језику *Ret-Fig* приказан је у раду „Ontology of Rhetorical Figures for Serbian“ (Младеновић and Mitrović, 2013). Онтологија *RetFig* описује 98 реторичких фигура. За сваку од њих дефинисано је којој реторичкој и којој лингвистичкој групи припада. Дефинисани су лингвистички опсег, објекти и елементи који учествују у креирању дате реторичке фигуре, међусобни однос објеката и елемената као и лингвистичке операције које учествују у процесу креирања реторичке фигуре. Све фигуре подељене су у четири групе: фигуре наглашавања (дикције), фигуре конструкције, фигуре речи (тропи) и фигуре проширења мисли. Фигуре наглашавања се заснивају на дејству одређених гласова у говору. Понављање одређених гласова или група гласова у говору, односно слова или група слова у тексту, њихово изостављање или уметање на неочекиваним местима, опонашање одређених звукова и шума из природе утиче на појачавање или смањење значаја језичких структура над којима се изводе. Ова врста фигура нема утицаја на значење структура над којима се граде и не мења их, већ само наглашава њихово основно значење. Фигуре конструкције настају мењањем реда речи у реченици или у некој другој већој целини (одломку, стиху и сл.) у односу на уобичајени тј. подразумевани, па се могу посматрати као фигуре поретка или реда. И ова врста фигура не мења основно значење језичких структура над којима се граде. Насупрот овим групама, фигуре речи (тропи) имају задатак да промене основно значење неке речи. Фигурама проширења мисли такође се мења основно значење језичке структуре која је комплекснија од речи.

Аутоматско препознавање реторичких фигура наглашавања (дикције) и конструкције у тексту могуће је спровести применом формализованих образаца како је то примењено у радовима (Gawryjolek et al., 2009; Hromada, 2011). На пример, *антиметабола* је фигура конструкције коју карактерише понављање речи или фраза из првог дела реченице и у другом, али обрнутим редоследом². У раду (Gawryjolek et al., 2009) описан је модел препознавања као и одговарајући алат за аутоматску анотацију (Java Annotation Tool Of Rhetoric — JANTOR) у коме се за *антиметаболу* дефинише синтаксички образац облика $[W_a] \dots [W_b] \dots [W_b] \dots [W_a]$ и којим се проналази појава ове фигуре у тексту, где су W_a and W_b речи или фразе које се у посматраној реченици у другом делу реченице појављују у инверзном поретку у односу на први. Резултати примене JANTOR-а показују да се реторичке фигуре из група фигура конструкције и наглашавања успешно проналазе на овај начин, јер је могуће дефинисати регуларне изразе који их једнозначно одређују. Обе ове класе фигура засноване су на синтаксичким или морфолошким операцијама у склопу граматичких

² Нпр. „Једемо да живимо, а не живимо да једемо.“; „Пиши као што говориш, говори као што пишеш.“

правила природног језика. Међутим структура реторичких фигура у класама фигура тропи и проширења мисли је таква да оне немају једнозначан образац примене на нивоу синтаксе.

У процесима аутоматског препознавања и анотације реторичких фигура које утичу на промену значења текста успешно се користе методе машинског учења. У радовима који су се бавили препознавањем *метонимије* коришћене су методе надгледаног машинског учења као што су листе одлучивања (Markert and Nissim, 2002), максимална ентропија (Farkas et al., 2007), к-најближих суседа (Leveling, 2007)³, стабла одлучивања, логистичка регресија (Nicolae et al., 2007) и др. Када је реч о *иронији* и *сарказму*, у раду (Tsur et al., 2010) коришћена је метода полу-надгледаног учења ради препознавања и класификације саркастичних твитова и коментара са сајта за електронску трговину Амазон. Метода потпорних вектора и логистичка регресија коришћене су у процесу класификације твитова на саркастичне и позитивно и негативно поларисане несаркастичне твитове у (González-Ibáñez et al., 2011).

С друге стране, показано је да се *иронија* и *сарказам* могу успешно детектовати и другим техникама. У (Carvalho et al., 2009) генерисно је осам образаца (као нпр. $P_{laugh} = (LOL|AH|EMO+)$, $P_{quote} = (ADJ_{pos}|N_{pos})\{1, 2\}$) који представљају структуралне индикаторе постојања *ироније* у коментарима објављиваним на новинским порталима на португалском језику. Метода групне расподеле рада (енг. crowd sourcing) у процесу детекције *ироније* и *сарказма* у приказима производа на сајту Амазон примењена је у (Filatova, 2012), док је у (Veale, 2012) дефинисана аутоматска метода екстракције семантичког знања из примера примене фигуре компарације ради препознавања *ироније*. У (Mitrović et al., 2015; Mladenović et al., 2016a) коришћена је метода групне расподеле рада у процесу детекције *компарације* у анотираном корпусу и предложен је алгоритам за аутоматску екстракцију детектованих кандидата на основу фреквентности појављивања у корпусу.

У овом раду у поступку детекције реторичких фигура из групе фигура речи (тропи) користи се метода онтолошког закључивања над онтологијом проистеклом из семантичке мреже Српски ворднет (SWN)⁴.

3. Онтолошко препознавање фигура речи (тропи)

Онтологија је један од видова репрезентације знања. Уколико је онтологија дата на неком формалном језику и складишти се у формату који је рачунарски

³ Коришћен је софтверски пакет TiMBL заснован на имплементацији алгоритма Memory-Based Learning који је проистекао из методе „k-nearest neighbor“.

⁴ <http://sm.jerteh.rs/MemberZone/eW3.aspx?id=miljanam>

читљив, реч је о формалној онтологији. Схваћена као „формална спецификација дељене концептуализације“ (Gruber, 1993), онтологија указује на онај вид знања који је могуће пренети, разменити и употребити. Према (Devedžić, 2010), главна сврха онтологија састоји се у дељењу и вишеструкој употреби знања од стране различитих интелигентних агената и апликација. Зависно од тога који део стварности описују, онтологије могу бити:

- онтологије највишег нивоа (енг. top level ontologies) — када моделирају опште концепте, а знање које репрезентују је свеобухватно, систематизовано и применљиво у великом броју апликација;
- доменске онтологије (енг. domain ontologies) — када се знања која репрезентују тичу једног домена или класе проблема;
- онтологије задатака (енг. task ontologies) или апликацијске онтологије (енг. application ontologies) — садрже само она знања неопходна за извршавање дате класе задатака.

У овом раду предлажемо методу идентификације реторичких фигура из класе фигура речи (тропи) коју, по дефиницији, чине оне фигуре чија се улога заснива на промени основног значења речи или фразе унутар једне језичке целине која може бити: реч, фраза, стих или реченица. Реторичке фигуре из ове класе своју улогу заснивају на:

- замени значења посматране речи тако што се она замени неком другом речи или фразом (нпр. фигура *иронија* се у примеру „*Красан си ми ти пријатељ*“ састоји у замени речи *лош* речју *красан* у посматраној реченици);
- додели вишеструког значења посматраној речи или фрази (нпр. фигура *синегдоха* где у примеру „*Добио је кров над главом*“, фраза *кров над главом* може добити и значење речи *сигурност*, *уточиште* и њихових синонима);
- генерисању новог значења посматране речи или фразе (фигура *оксиморон* која спаја појмове супротних значења у нови појам, нпр. *виртуелна стварност*).

У овим случајевима није могуће дефинисати морфолошко-синтаксичке шаблоне, али је могуће утврдити какве семантичке релације постоје између учесника који граде одређену фигуру и на основу тога дефинисати правила закључивања која се могу применити над онтологијом каква је проистекла из семантичке мреже ворднет (Fellbaum, 1998). Фигуре које припадају класи тропи добили смо SPARQL упитом над доменском онтологијом *RetFig* (Mladenović and Mitrović, 2013).

```
select distinct ?figura
where { ?figura ont:jeNadObjektom ?objekt .
```

```
?figura ont:jeRetorickaGrupa ?retGrupa.
?objekt ont:naziv ?nazivObjekta.
?retGrupa ont:naziv ?nazivRetGrupe.
FILTER (?nazivObjekta="REC" &&
?nazivRetGrupe="FIGURE_TROPI" )}
order by ?figura
```

Тако је добијен скуп од 26 реторичких фигура (слика 1) које имају заједничку особину да припадају групи фигура тропи (?nazivRetGrupe="FIGURE_TROPI") и које карактерише неки вид промене уобичајеног значења једне речи (?nazivObjekta="REC"). Класа на врху овог дела онтологије *RetFig* означена је као Tropi. Она садржи три подкласе (VišestrukoZnačenje, IzmenjenoZnačenje, NovoZnačenje) којима се дефинише улога појединачних фигура, односно вид промене уобичајеног значења речи над којом се фигура примењује. Идеја је да се, уколико је могуће, свака од тих фигура дефинише применом правила (изражених помоћу Semantic Web Rule Language — SWRL)⁵ над онтологијом SWN.

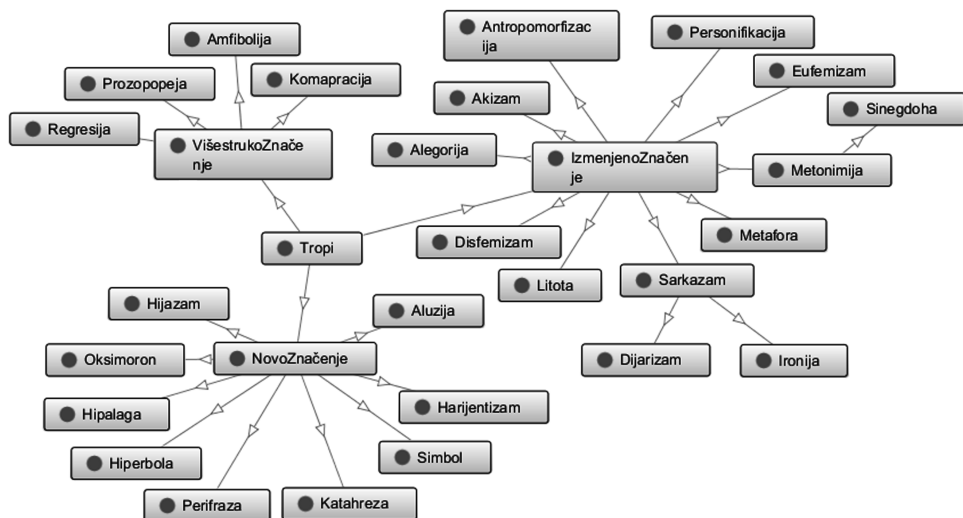
На примеру фигуре *иронија* (енг. irony) биће објашњен начин изградње правила над онтологијом SWN помоћу којих се може открити појава одговарајуће фигуре у тексту. Вербална иронија је појава речи или израза којим се једно каже, а супротно мисли и разуме. Право значење речи прикривено је и супротно исказаном значењу израза. Примери вербалне ироније заснивају се на неколико врста семантичких релација (Bagić, 2012), а често коришћени облици у српском језику заснивају се на употреби придева коме се даје значење њему подразумеваног супротног (антонимског) придева као у примерима:

- „Он је баш генијалан!“ (Скривено значење је тврдња да је неко заправо глуп.)
- „Види се да је жгољав!“ (Скривена тврдња да је неко дебео.)

Друга форма се састоји од везе именице и придева, али се придеву додељује скривено значење оног придева који се уобичајено везује уз наведену именицу.

- „Брз је као корњача.“
- „Храбар је као зец!“

⁵ Semantic Web Rule Language — SWRL је језик који се користи у области семантичког веба за представљање формалних логичких израза, а добијен је комбиновањем особина језика OWL DL (Web Ontology Language for Description Logic) и језика RML (Rule Markup Language).



Слика 1: Део таксономије онтологије изведене из онтологије *RetFig* који се односи на класу фигура речи (тропи)

У раду (Veale and Hao, 2009) користи се назив „иронијна компарација“ како би се описала специфична врста ироније изведена из компарације. У наведеним примерима придев *брз* добија скривено значење придева који представља природну карактеристику именице *корњача*, а то је придев *спор*, док придев *храбар* добија скривено значење придева који је природна карактеристика именице *зеи*, а то је придев *плашљив*. Семантичке релације које постоје у онтологији SWN и које се могу користити при генерисању кандидата за детекцију фигуре *иронија* су релације *specificOf/specifiedBy* и релација *near_antonym*. Релације *specificOf/specifiedBy* представљају пар инверзних релација (Mladenović et al., 2016a) којима се повезују инстанце класе именичких синсетова са инстанцама синсетова њима специфичних придева. Релација *near_antonym* међусобно повезује инстанце класа синсетова придева који су директни антоними. Стога се правило за генерисање иронијних релација, примењено над онтологијом SWN, може изразити као OWL правило

$$\{?n : \text{specifiedBy } ?p1. ?p1 : \text{near_antonym } ?p2\} \Rightarrow \{?n : \text{Irony } ?p2\}$$

или у форми SWRL правила

$$\text{specifiedBy}(?n, ?p1), \text{near_antonym}(?p1, ?p2) \rightarrow \text{Irony}(?n, ?p2) \quad (1)$$

Правило (1) може се добити и SPARQL упитом (слика 2) над онтологијом SWN.

SPARQL query:

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX swn30: <http://www.mmijana.com/swn30#>
SELECT ?adjective1 ?adjective2 ?noun
WHERE {
 ?adjective1 swn30:near_antonym ?adjective2.
 ?adjective2 swn30:specificOf ?noun.
}
order by ?adjective1

adjective1	adjective2	noun
hrabar	plašljiv	zec
brz	spor	kornjača
brz	spor	puž
spor	brz	strela
spor	brz	ideja
lak	težak	slon
vruć	hladan	kamen

Слика 2: Кандидати за генерисање фигура *иронија* и *компарација*

Резултат упита даје кандидате фигуре *иронија* у форми пара (adjective2, noun), нпр. (*храбар, зец*), али даје и кандидате фигуре *компарација* (енг. simile) које чине парови (adjective1, noun), нпр. (*плашљив, зец*), што је еквивалентно OWL правилу

$$\{?n : specifiedBy ?p1\} \Rightarrow \{?n : Simile ?p1\}$$

или

$$specifiedBy(?n, ?p1) \rightarrow Simile(?n, ?p1) \tag{2}$$

Резултат упита са слике 2 аналоган је резултату закључивања на основу правила (1), али се из упита може уочити да у природном језику није уобичајено коришћење неких од кандидата за генерисање фигуре *иронија*. На основу интуиције која се у овом тренутку не доказује, може се рећи да су природни кандидати фигуре *иронија* они код којих је поларитет осећања субјекта (?adjective1) у RDF тројци⁶ ?adjective1 swn30:near_antonym ?adjective2 — позитиван.

⁶ RDF је модел података који користи форму тројке или триплета „субјекат-предикат-објекат“ (subject–predicate–object) да би описао ресурсе семантичког веба, обезбедио

На пример, природнији примери ироније исказани су тврдњама: „*Брз је као корњача*“ или „*Лак је као слон*“ него „*Спор је као стрела*“, или „*Тежак је као перце*“. У том смислу, претрага се може унапредити додатним условом који узима у обзир оне RDF тројке у којима је вредност негативног поларитета осећања субјекта једнака нули. Како сваки синсет у семантичкој мрежи Српски ворднет има етикету (Mladenović and Mitrović, 2014), као и вредност, позитивног и негативног сентимент поларитета, што је добијено успостављањем једнозначних релација са ресурсом SentiWordNet (Baccianella et al., 2010), то се, у циљу идентификације поларитета осећања субјекта у RDF тројци, упит може проширити контролом својства податка (dataproperty) `swn30:sentimentNegative`, а на основу вредности тог својства бирају се (филтрирају) само они елементи код којих је вредност тог својства једнака нули (слика 3).

SPARQL query:		
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>		
PREFIX owl: <http://www.w3.org/2002/07/owl#>		
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>		
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>		
PREFIX swn30: <http://www.mmijana.com/swn30#>		
SELECT ?adjective1 ?adjective2 ?noun		
WHERE {		
?adjective1 swn30:near_antonym ?adjective2.		
?adjective2 swn30:specificOf ?noun.		
?adjective1 swn30:sentimentNegative ?sent.		
FILTER regex(str(?sent), "0,00000")		
}		
ORDER by ?adjective1		
adjective1	adjective2	noun
hrabar	plašljiv	zec
brz	spor	kornjača
brz	spor	puž
vruć	hladan	kamen

Слика 3: Побољшање претраге кандидата за генерисање фигуре *иронија*

Одговарајуће SWRL правило, које укључује додатни услов приказан на слици 3, дато је у наставку

$$\begin{aligned}
 &specifiedBy(?n, ?p1), near_antonym(?p1, ?p2), \\
 &sentiment_negative(?p1, ?sent), swrlb : equal(?sent, 0) \rightarrow Irony(?n, ?p2)
 \end{aligned}
 \quad (3)$$

Размотримо сада и структуру реторичке фигуре *оксиморон* која, по дефиницији, представља спајање појмова супротних значења у нови појам и чији

складиштење података у графовским базама података и представљање знања у онтолошким моделима.

су примери у SWN бази: *виртуелна ствараност*, *гласна тишина*, *луда памет*, *живи фосил*, *ватрени лед*, и др. Правила за генерисање фигуре *оксиморон* примењена над онтологијом SWN могу се изразити као:

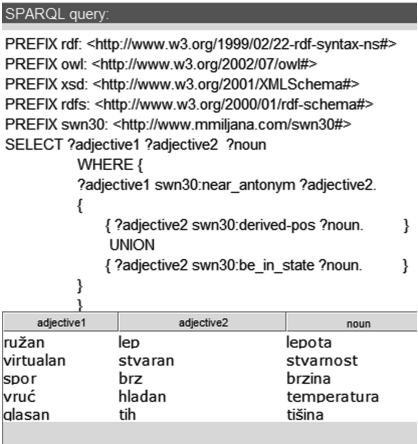
$$\begin{aligned} \{?p1 : near_antonym\ ?p2.\ ?p2 : derived-pos\ ?n\} &\Rightarrow \{?p1 : oxymoron\ ?n\} \\ \{?p1 : near_antonym\ ?p2.\ ?p2 : be_in_state\ ?n\} &\Rightarrow \{?p1 : oxymoron\ ?n\} \end{aligned}$$

или у форми SWRL правила:

$$near_antonym(?p1, ?p2), derived-pos(?p2, ?n) \rightarrow oxymoron(?p1, ?n) \tag{4}$$

$$near_antonym(?p1, ?p2), be_in_state(?p2, ?n) \rightarrow oxymoron(?p1, ?n) \tag{5}$$

Правила (4) и (5) могу се добити и SPARQL упитом (слика 4) над онтологијом SWN. Резултат упита даје кандидате за генерисање фигуре *оксиморон* у форми пара (adjective1, noun), нпр. (*гласан*, *тишина*). Са слике 4 може се уочити да RDF тројке (adjective1, adjective2, noun) које садрже антонимске придеве постају кандидати за инстанцирање ове фигуре, када задовољавају услов постојања релација *derived-pos* или *be_in_state*, које представљају лексичке релације између синсетова различитих врста речи (енгл. cross-part of speech relations — XPoS), у овом случају између придева и именица (Koeva et al., 2008).



Слика 4: Кандидати за генерисање фигуре *оксиморон*

Поред лексичко-семантичких релација које могу бити индикатор постојања одређених тропи фигура, може се користити и имплицитна релација синонимије којом се могу проналазити кандидати за инстанцирање фигуре *перифраза* (Mitrović, 2014), фигуре којом се један појам описује или замењује са више речи коришћењем неких битних особина тог појма (слика 5). Тако се реч *Pariz* може заменити синонимима из ворднета: *Grad svetlosti*, *Prestonica Francuske*, *Glavni grad Francuske*⁷, па парови попут (*Pariz*, *Grad svetlosti*) постају кандидати за инстанце ове фигуре.

SPARQL query:		
<pre> PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> PREFIX owl: <http://www.w3.org/2002/07/owl#> PREFIX xsd: <http://www.w3.org/2001/XMLSchema#> PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#> PREFIX swn30: <http://www.mmiljana.com/swn30#> SELECT ?noun ?synonymNoun WHERE { ?noun swn30:hasNounWord ?synonym. ?noun rdfs:label ?os. ?synonym swn30:literal ?synonymNoun FILTER (CONTAINS (UCASE(str(?os)), "PARIZ")) } </pre>		
noun	synonymNoun	
Pariz	"prestonica Francuske"	
Pariz	"glavni grad Francuske"	
Pariz	"Grad svetlosti"	
Pariz	"Pariz"	

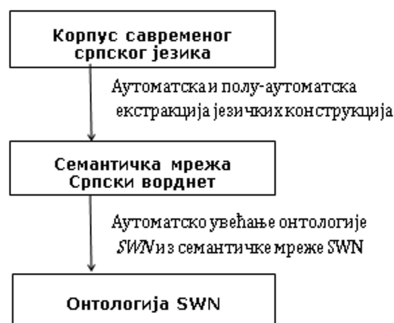
Слика 5: Кандидати за генерисање фигуре *перифраза*

4. Учење онтологије SWN

У претходном одељку описана су правила, дефинисана у онтологији SWN, којима се аутоматски генеришу инстанце неких реторичких фигура које припадају групи фигура тропи. Међутим, главни проблем у инжењерингу једне онтологије према (Devedžić, 2010) је њено оспособљавање да се мења истовремено са увећањем и изменом доменских знања. Способност онтологије да непрестано „учи“ (енг. *ontology learning*) тако што се аутоматским и полу-аутоматским поступцима екстрахују и аотирају нова знања из посматраног домена и интегришу са постојећим знањима у онтологији, у многоме зависи

⁷ Сви примери екстраховани из ворднета приказани су у свом оригиналном облику, односно на латиници.

од алата који се у ту сврху примењују. Како је задатак онтологије SWN да препознаје различите реторичке фигуре, врло је важно да она „учи“ различите облике фигуративног говора из примера на природном језику. С обзиром да се процес тог учења заснива на препознавању учесника у свим семантичким релацијама онтологије SWN, а посебно у релацијама као што су: *specificOf*, *specifiedBy*, *near_antonym*, *derived-pos*, *be_in_state* и *synonym*), уколико се обезбеди аутоматско и интензивно ручно и континуирано проширење онтологије SWN овим релацијама и ентитетима који се повезују овим релацијама, добиће се онтологија која ће бити потпунија и самим тим ефикаснија у решавању постављених задатака. Дијаграм учења онтологије SWN приказан је на слици 6.



Слика 6: Учење онтологије SWN из Корпуса савременог српског језика

Учење онтологије SWN засновано на аутоматској методи проширења семантичке мреже Српски ворднет (Mladenović et al., 2016a) користи анотирани део електронског (дигиталног) корпуса савременог српског језика⁸ (Utvić, 2014) ради генерисања кандидата за проширење релацијама *specificOf* и *specifiedBy*. У овом раду се предлаже полу-аутоматска метода проширења семантичке мреже Српски ворднет заснована на употреби веб алата којим се олакшава повезивање синсетова помоћу неке од лексичко-семантичких релација дефинисаних у SWN. Корисничко радно окружење приказано је на слици 7 и представља алат којим се повезују два жељена синсета једном од релација које су на располагању. На левој страни слике 7, за унуту реч (нпр. *lep*) приказују се, испод ње, сва значења ове речи која су дефинисана у Српском ворднету. На десној страни исте слике, за унуту реч (нпр. *primerRecsslika*), такође испод, приказују се сва њена значења дефинисана у

⁸ <http://korpus.matf.bg.ac.rs/prezentacija/korpus.html>

Српском ворднету. Корисник бира одговарајуће значење сваке од унетих речи и повезује их лексичко-семантичком релацијом коју бира из скупа релација датих у средњој колони. Уколико корисник одабере релацију за коју постоји инверзна релација (нпр. *SpecificOf-SpecifiedBy*), аутоматски се генерише и одговарајућа инверзна релација између синсетова концепата који су одабрани на основу значења.

Prikazi

Poveži relacijom

Prikazi

Synset	Literal	Sense	Definition		1 2	Synset	Literal	Sense	def
ENG30-00217728-a	lep	1	Koji ushićuje čula, uzbuđuje duh ili izaziva emocionalno divljenje.	☒	Relation	ENG30-13937075-n	slika	4v	situacija koja se tretira kao osmotriv predmet
					holo_portion				
					TopicDomain				
					causes	ENG30-03876519-n	slika	1	Grafička umetnost koja se sastoji od umetničke kompozicije dobijene nanošenjem boje na neku površinu.
					hypernym				
					holo_part				
					also_see	ENG30-03931044-n	slika	2	Vizuelna reprezentacija objekta, scene, osobe ili apstrakcije, proizvedena na nekoj površini.
					TopicDomainMember				
					be_in_state	ENG30-03314028-n	slika	x	Jedna od dvanaest karata iz špila na čijem je licu slika.
					Entailment	ENG30-07201804-n	slika	4a	Grafički ili živ verbalni opis.
					holo_member	ENG30-14513489-n	slika	4ax	Okrugljenje u kome se odvija priča ili dramska radnja.
					Hyponym				
					substanceMeronym				
					specifiedBy				
					derived-vn				

Слика 7: Веб алат за полу-аутоматско проширење семантичке мреже Српски ворднет; у овом случају повезани су придев *lep* и именица *слика* (slika) паром међусобно инверзних релација *SpecificOf / SpecifiedBy*

5. Евалуација методе онтолошког препознавања реторичке фигуре *компарација*

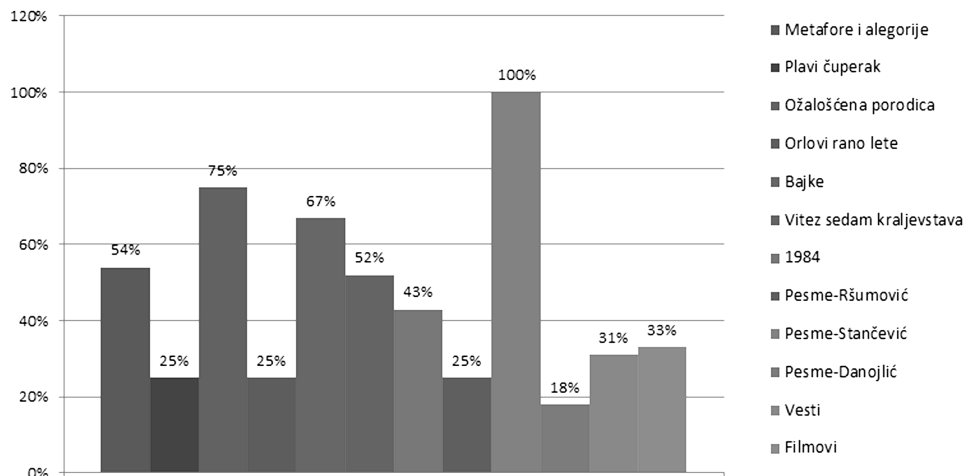
У циљу оцене степена препознавања реторичке фигуре *компарација*, у текстовима на српском језику, спроведен је тест изведен над колекцијом различитих текстова која се састоји од 10 дигитализованих дела разноврсног жанра: дечје песме, бајке, комедије, романи и студије наведених у додатку А. Овом скупу придружени су и текстови који су у раду (Mladenović et al., 2016b) коришћени као скупови за тестирање класификатора текстова на основу осећања, а који су састављени од оцена филмова и вести са новинских портала на српском језику (ступци означени лабелама *vesti* и *filmovi* на слици 8). Изграђен је алат којим се, уз помоћ регуларних израза, улазни текст парсира

тако да се екстрахују изрази у форми кандидата за препознавање фигуре *компарација*. Регуларним изразима из наведених текстова екстраховане су структуре облика „РЕЧ као РЕЧ {1,2}“, „РЕЧ *попут* РЕЧ {1,2}“. Из екстрахованих структура, ручно су уклоњене све оне које нису облика „РЕЧ као ИМЕНИЦА“, „РЕЧ као ПРИДЕВ ИМЕНИЦА“, „РЕЧ *попут* ПРИДЕВ ИМЕНИЦА“. Најзад, изграђен је и алат који, на основу правила (2) (в. одељак 3.) у онтологији SWN закључује да ли дата структура представља фигуру компарација.

Језичке структуре које су добијене овим тестом указују на то да се у поетским текстовима могу наћи поређења неувобичајена у свакодневном говору. На пример, у песмама Данојлића, се јављају структуре *добар као слон, спокојан као крава, устрептала као стрела* и сл. У песмама Ршумовића могу се наћи: *плаво као ћилим, зелено као тепих, ружан као четка*, итд. На крају овог теста, анализирана је структура скупова за тестирање *filmovi* и *vesti* којима је тестиран SAFOS — систем за анализу осећања текстова на српском језику (Mladenović et al., 2016b), како бисмо оценили могућност за његову даљу доградњу у смислу препознавања фигуративног говора. Резултати су показали ниску, али уједначену тачност⁹ препознавања од око једне трећине укупног броја појављивања фигуре *компарација*. На примерима структура које су екстраховане из ових скупова као што су: *okarakterisan kao triler, poznat kao zemlja, urađen kao nastavak* и сл. може се видети да би у текстовима било могуће детектовати и лажно позитивне кандидате за инстанцирање фигуре *компарација*. Међутим, како је препознавање фигуре вршено помоћу знања у онтологији SWN, постигнута је потпуна прецизност, чиме су лажно позитивни кандидати одбачени¹⁰. С друге стране, непрепознате структуре облика: *oštrim poput brijača, hrabrim kao SUPERMEN, puknuti kao vidik, razmnožavati se kao vinska mušica* указују на низак ниво одзива класификације и потребу испитивања и другачијих језичких структура од описаних, узимањем у обзир и других врста сем именица (личних имена, израза, скраћеница и др.) са једне, као и глагола са друге стране језичке структуре, из којих би онтологија SWN могла учити. Резултати тачности класификације фигуре *компарација*, изведене над скупом улазних текстова који се састоји од 10 дигитализованих дела и два скупа текстова којим је тестиран класификатор текстова на основу осећања, приказани су на слици 8.

⁹ Тачност (енг. *assiguacy*) представља једну од мера оцене квалитета класификације и представља однос збира тачно позитивно и тачно негативно класификованих примерака и збира свих примерака у тестираној колекцији.

¹⁰ У онтологији SWN не постоји релација *specifiedBy* између именице и придева, осим уколико то није придев верификован методама учења онтологије SWN (в. одељак 4.).



Слика 8: Тачност препознавања фигуре *компарација* онтологијом SWN у текстовима на српском језику

Евалуација система приказана је на поступку препознавања фигуре *компарација* у роману „Витез седам краљевстава“ Џорџа Мартина, с обзиром да се ради о најобимнијем интегралном тексту у скупу тестираних текстова. Применом регуларних израза пронађено је укупно 135 конструкција облика „РЕЧ (као|попут) РЕЧ{1,2}“. Од тога је 107 конструкција које представљају фигуру *компарација*, а 28 није носило то значење као у примерима: „*Tu sam pecaо kao dečак*“, „*Lim je imao naoštren kolac koji bi mogao poslužiti kao kop-lje*“, „*Dve godine je služio kao раж*“ и сл. Ова класификација је извршена ручно, а затим је такав скуп подвргнут онтолошкој класификацији. То значи да је за сваки од 135 парова (реч, именица) добијених из датих језичких структура проверено да ли задовољавају правило (2), тј. да ли у посматраном уређеном пару реч представља придев, а затим да ли је уређени пар придев-именица повезан релацијом *specificOf*. Неки примери препознатих фигура компарације су: „*Bele kao kost bile su put i kosa Brindena Rečnog*“..., „*Bila su hladna kao ka-men, ali ih je bilo divno videti*“, „... *bilo je crveno kao krv*“, а резултати онтолошке класификације приказани су у Табели 1, матрицом конфузије.

На основу матрице конфузије, параметри оцене система дати су у табели 2, а статистичка оцена класификатора представљена је ROC кривом на слици 9, где је добијено да је $AUC > 0,5$ за интервал поверења 95%.

класификација фигуре компарација		ручна класификација	
		да	не
оцена класификатора	да	$tp = 42$	$fp = 0$
	не	$fn = 65$	$tn = 28$

Табела 1: Матрица конфузије система класификације језичких структура које су екстраховане из романа „Витез седам краљевстава“ Џорџа Мартина на оне које представљају фигуру *компарација* и оне које то нису.

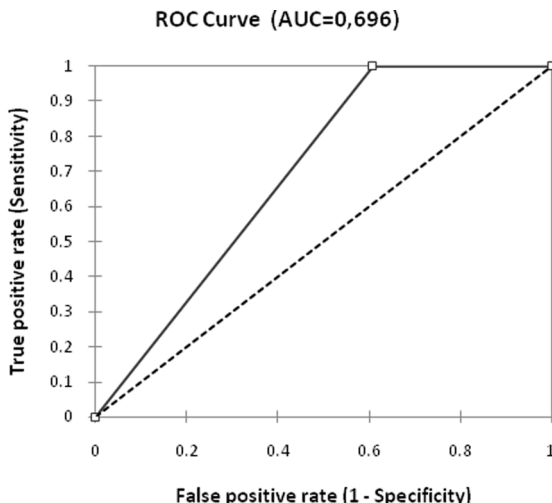
параметри оцене система класификације фигуре <i>компарација</i>	Прецизност = $\frac{tp}{tp+fp}$	Одзив = $\frac{tp}{tp+fn}$	$F_1 = \frac{2PR}{P+R}$	Тачност
Мере	1,000	0,393	0,564	0,518

Табела 2: Параметри оцене система класификације фигуре *компарација* у роману „Витез седам краљевстава“ Џорџа Мартина

Receiver operating characteristic curve (ROC) је једна од метода евалуације класификатора која користи графички приказ односа сензитивности ($\frac{tp}{tp+fn}$) и специфичности ($\frac{tn}{tn+fp}$) за сваки могући резултат на тесту. На слици 9 приказана је (пуном линијом) ROC крива класификатора фигуре *компарација*, док се дијагонална линија (приказана испрекиданом линијом) назива дијагоналом случајног исхода и представља исходе случајне класификације. Површина испод ROC криве (енг. Area Under a Curve — *AUC*) интерпретира се као вероватноћа да случајно одабран позитивно означен примерак фигуре *компарација* има већи скор него примерак означен негативно. *AUC* представља меру тачности класификатора и може имати вредности од $AUC = 0,5$ (када се ROC крива поклапа са дијагоналом случајног исхода) до $AUC = 1$ (у случају апсолутног раздвајања класа). Класификација фигуре *компарација* постигла је вредност $AUC = 0,696$.

Анализирањем резултата класификације, могу се уочити разлози ниског одзива. Постоје три групе структура које нису препознате предложеним системом:

1. структуре које садрже личну именицу као у примерима: „*pametan kao princ Eris*“, „*visok kao Dank*“, и сл.;
2. структуре које не садрже придев и именицу, већ глагол и именицу као у примерима: „*Onda se okrenu kao vikor, i jurnu u noć*“ и „*Jaje ulete kao bez duše*“;



Слика9: Статистичка оцена система класификације фигуре *компарација* у роману „Витез седам краљевстава“ Џорџа Мартина

3. структуре које садрже придев и именицу, али је поредак такав да није могуће једнозначно детектовати фигуру на основу структура „ПРИДЕВ као/попут ИМЕНИЦА“ и „ПРИДЕВ као/попут ПРИДЕВ ИМЕНИЦА“, већ би се посматрани оквир морао проширити. Примери таквих структура су пронађени у реченицама: „... *gole kao od majke rođene, brašnjave od glave do pete*“, „... *čije su vode svetlucale crveno i zlatno, sjajne kao ploča kovanog bakra*“, „*Zbog blede kože i kao kreč bele kose ličio je na živi leš.*“

Структуре које садрже личну именицу могу се препознавати укључивањем алата за препознавање именованих ентитета и обрасца претраге „ПРИДЕВ као ВЛАСТИТА_ИМЕНИЦА“. Оне које садрже глагол и именицу могу се идентификовати употребом алата за означавање врстама речи и обрасца „ГЛАГОЛ као ИМЕНИЦА“, а оне измењеног поретка могу се препознавати обрасцима који то узимају у обзир, нпр. „као ИМЕНИЦА [ВЕЗНИК|ПРЕДЛОГ]? ПРИДЕВ“.

6. Закључак

У овом раду предложена је метода препознавања реторичких фигура из групе фигура речи (тропи) које се не могу препознавати на основу синтаксичко-

морфолошких образаца. Метода користи правила дефинисана у онтологији SWN, како би препознала да ли језичка структура екстрахована из текста задовољава постављено правило и на тај начин може бити означена као одговарајућа реторичка фигура. Метода је тестирана откривањем реторичке фигуре *компарација* над скупом улазних текстова који се састоји од 10 дигитализованих дела и два скупа текстова којима је претходно тестиран класификатор текстова на основу осећања, а статистичка оцена система спроведена је над текстом романа „Витез седам краљевстава“ Џорџа Мартина. Резултати статистичке оцене показали су да класификација овом методом ($AUC = 0,696$) у случају реторичке фигуре *компарација* показује осредње резултате, али прецизност класификације од 100% указује на то да се даљим учењем онтологије, на основу предложене полуаутоматске методе, могу побољшати резултате класификације ове фигуре.

У даљем раду, када се ради о фигури *компарација*, биће уведени обрасци описани у одељку 5. ради препознавања структура које нису препознате предложеним системом. Када је реч о другим фигурама, радиће се на изградњи скупова за тестирање класификатора реторичких фигура *иронија*, *сарказам* и *оксиморон*, како би позитивно класификоване структуре могле бити коришћене у процесу изградње предиктора класификатора текста на основу осећања. Радиће се и на проширењу скупа онтолошких правила тако да укључују поступке генерализације и спецификације концепата који већ задовољавају правила постављена овим системом.

Литература

- Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In Calzolari et al. (2010). ISBN 2-9517408-6-7. URL <http://www.lrec-conf.org/proceedings/lrec2010/summaries/769.html>.
- Krešimir Bagić. *Rječnik stilskih figura*. Školska knjiga, Zagreb, 2012.
- Francesco Barbieri, Francesco Ronzano, and Horacio Saggion. UPF-taln: SemEval 2015 Tasks 10 and 11. Sentiment Analysis of Literal and Figurative Language in Twitter. In Daniel M. Cer, David Jurgens, Preslav Nakov, and Torsten Zesch, editors, *Proceedings of the 9th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2015, Denver, Colorado, USA, June 4-5, 2015*, pages 704–708. The Association for Computer Linguistics, 2015. ISBN 978-1-941643-40-2. URL <http://aclweb.org/anthology/S/S15/S15-2119.pdf>.
- Daniel G. Bobrow. A Question-Answering System for High School Algebra Word Problems. In *Proceedings of AFIPS Conference, 26. FJCC, Part I*, pages 591–614, New York, 1964. ACM. doi: 10.1145/1464052.1464108.

- Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner, and Daniel Tapias, editors. *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010, 17-23 May 2010, Valletta, Malta*, 2010. European Language Resources Association. ISBN 2-9517408-6-7. URL <http://www.lrec-conf.org/proceedings/lrec2010/index.html>.
- Paula Carvalho, Luís Sarmiento, Mário J. Silva, and Eugénio de Oliveira. Clues for Detecting Irony in User-generated Contents: Oh...!! It's "So Easy" ;-). In *Proceedings of the 1st International CIKM Workshop on Topic-sentiment Analysis for Mass Opinion*, pages 53–56, New York, 2009. ACM.
- Paula Carvalho, Luís Sarmiento, Jorge Teixeira, and Mário J. Silva. Liars and Saviors in a Sentiment Annotated Corpus of Comments to Political Debates. In Lin et al. (2011), pages 564–568. ISBN 978-1-932432-88-6. URL <http://www.aclweb.org/anthology/P11-2099>.
- Vladan Devedžić. *Semantic Web and Education*. Integrated Series in Information Systems. Springer-Verlag New York Inc., 2010. ISBN 9781441942012.
- Richard Farkas, Eszter Simon, Gyorgy Szarvas, and Daniel Varga. Gyder: Maxent Metonymy Resolution. In Eneko Agirre, Lluís Màrquez, and Richard Wicentowski, editors, *SemEval '07: Proceedings of the 4th International Workshop on Semantic Evaluations*, pages 161–164, Stroudsburg, PA, USA, 2007. Association for Computational Linguistics.
- Dan Fass. met*: A Method for Discriminating Metonymy and Metaphor by Computer. *Computational Linguistics*, 17(1):49–90, 1991. ISSN 0891-2017.
- Christiane Fellbaum. *WordNet: An Electronic Lexical Database*. Language, Speech, and Communication. A Bradford Book, 1998. ISBN 978-0262061971.
- Elena Filatova. Irony and Sarcasm: Corpus Generation and Analysis Using Crowdsourcing. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey, 2012. European Language Resources Association (ELRA). ISBN 978-2-9517408-7-7.
- Jakub Gawryjolek, Chrysanne Di Marco, and Randy A. Harris. An Annotation Tool for Automatically Detecting Rhetorical Figures — System Demonstration. In *Proceedings of the IJCAI-09 workshop on Computational Models of Natural Argument*, Pasadena, CA, 2009.
- Roberto I. González-Ibáñez, Smaranda Muresan, and Nina Wacholder. Identifying Sarcasm in Twitter: A Closer Look. In Lin et al. (2011), pages 581–586. ISBN 978-1-932432-88-6. URL <http://www.aclweb.org/anthology/P11-2102>.
- Thomas R. Gruber. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5(2):199–220, 1993. ISSN 1042-8143. doi: 10.1006/knac.

- 1993.1008. URL <http://dx.doi.org/10.1006/knac.1993.1008>. Special issue: Current Issues in Knowledge Modeling.
- Yanfen Hao and Tony Veale. An Ironic Fist in a Velvet Glove: Creative Misrepresentation in the Construction of Ironic Similes. *Minds and Machines*, 20 (4):635–650, 2010. doi: 10.1007/s11023-010-9211-1. URL <http://dx.doi.org/10.1007/s11023-010-9211-1>.
- Andrew Hardie, Veronika Koller, Paul Rayson, and Elena Semino. Exploiting a Semantic Annotation Tool for Metaphor Analysis. In Davies, Matthew and Rayson, Paul and Hunston, Susan and Danielsson, Pernilla, editor, *Proceedings of the Corpus Linguistics 2007 Conference*, Birmingham, 2007. University of Birmingham.
- Randy Harris and Chrysanne Di Marco. Constructing a Rhetorical Figuration Ontology. In *Symposium on Persuasive Technology and Digital Behavior Intervention, Convention of the Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB)*, pages 47–52, Edinburgh, 2009.
- Daniel Devatman Hromada. Initial Experiments with Multilingual Extraction of Rhetoric Figures by means of PERL-compatible Regular Expressions. In Irina P. Temnikova, Ivelina Nikolova, and Natalia Konstantinova, editors, *Proceedings of the Student Research Workshop associated with The 8th International Conference on Recent Advances in Natural Language Processing, RANLP 2011, 13 September, 2011, Hissar, Bulgaria*, pages 85–90. RANLP 2011 Organising Committee, 2011. URL <http://www.aclweb.org/anthology/R11-2013>.
- Ashley R. Kelly, Nike A. Abbott, Randy Allen Harris, Chrysanne DiMarco, and David R. Cheriton. Toward an Ontology of Rhetorical Figures. In Júnia Coutinho Anacleto, Renata Pontin de Mattos Fortes, and Carlos J. Costa, editors, *Proceedings of the 28th Annual International Conference on Design of Communication, SIGDOC 2010, São Carlos, São Paulo state, Brazil, September 26-29, 2010*, pages 123–130. ACM, 2010. ISBN 978-1-4503-0403-0. doi: 10.1145/1878450.1878471. URL <http://doi.acm.org/10.1145/1878450.1878471>.
- Alistair Kennedy and Diana Inkpen. Sentiment Classification of Movie Reviews Using Contextual Valence Shifters. *Computational Intelligence*, 22(2):110–125, 2006. doi: 10.1111/j.1467-8640.2006.00277.x. URL <http://dx.doi.org/10.1111/j.1467-8640.2006.00277.x>.
- Svetla Koeva, Cvetana Krstev, and Duško Vitas. Morpho-Semantic Relations in WordNet - a Case Study for two Slavic Languages. In *Proceedings of Global Word-Net Conference*, pages 239–253, Szeged, 2008. University of Szeged, Department of Informatics.
- Veronika Koller, Andrew Hardie, Paul Rayson, and Elena Semino. Using a Semantic Annotation Tool for the Analysis of Metaphor in Discourse. *Metaphorik.de*, 15: 141–160, 2008.
- Johannes Leveling. FUH (FernUniversität in Hagen): Metonymy Recognition Using Different Kinds of Context for a Memory-Based Learner. In Eneko Agirre, Lluís

- Màrquez, and Richard Wicentowski, editors, *SemEval '07: Proceedings of the 4th International Workshop on Semantic Evaluations*, pages 153–156, Stroudsburg, PA, USA, 2007. Association for Computational Linguistics.
- Dekang Lin, Yuji Matsumoto, and Rada Mihalcea, editors. *The 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference, 19-24 June, 2011, Portland, Oregon, USA - Short Papers*, volume 2, 2011. The Association for Computer Linguistics. ISBN 978-1-932432-88-6.
- Katja Markert and Malvina Nissim. Metonymy Resolution As a Classification Task. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10*, EMNLP '02, pages 204–213, Stroudsburg, PA, USA, 2002. Association for Computational Linguistics. doi: 10.3115/1118693.1118720. URL <http://dx.doi.org/10.3115/1118693.1118720>.
- Zachary J. Mason. CorMet: A Computational, Corpus-based Conventional Metaphor Extraction System. *Computational Linguistics*, 30(1):23–44, 2004. ISSN 0891-2017. doi: 10.1162/089120104773633376.
- Ruslan Mitkov. *Anaphora Resolution*. Longman, Cambridge, UK, 2002.
- Jelena Mitrović. Electronic Tools and Resources for Multi-Word Unit Detection and Research in Serbian. The 2th General Meeting of The IC1207 COST Action, PARSEME, Athens, Greece, 10–11 March, 2014, 2014. URL <http://typo.uni-konstanz.de/parseme/images/Meeting/2014-03-11-Athens-meeting/PosterAbstracts/Mitrovic.pdf>.
- Jelena Mitrović, Miljana Mladenović, and Cvetana Krstev. Adding MWEs to Serbian Lexical Resources Using Crowdsourcing. The 5th PARSEME General Meeting, Iași, Romania, 23–24 September 2015, 2015. URL <http://typo.uni-konstanz.de/parseme/images/Meeting/2015-09-23-Iasi-meeting/WG1-MITROVIC-MLADENOVIC-KRSTEV-poster.pdf>.
- Miljana Mladenović and Jelena Mitrović. Ontology of Rhetorical Figures for Serbian. In Ivan Habernal and Václav Matousek, editors, *Text, Speech, and Dialogue - 16th International Conference, TSD 2013, Pilsen, Czech Republic, September 1-5, 2013. Proceedings*, volume 8082 of *Lecture Notes in Computer Science*, pages 386–393. Springer, 2013. ISBN 978-3-642-40584-6. doi: 10.1007/978-3-642-40585-3_49. URL http://dx.doi.org/10.1007/978-3-642-40585-3_49.
- Miljana Mladenović and Jelena Mitrović. Semantic Networks for Serbian: New Functionalities of Developing and Maintaining a WordNet Tool. In Gordana Pavlović-Lazetić, Cvetana Krstev, Ivan Obradović, and Duško Vitas, editors, *Natural Language Processing for Serbian: Resources and Applications*, pages 1–11, Belgrade, 2014. University of Belgrade, Faculty of Mathematics. ISBN 978-86-7589-088-1.
- Miljana Mladenović, Jelena Mitrović, and Cvetana Krstev. A Language-Independent Model for Introducing a New Semantic Relation between Adjectives and Nouns

- in a WordNet. In Verginica Barbu Mititelu, Corina Forăscu, Christiane Fellbaum, and Piek Vossen, editors, *The Proceedings of Eighth Global WordNet Conference 2016*, pages 218–225, Bucharest, Romania, 2016a. ISBN 978-606-714-239-6.
- Miljana Mladenović, Jelena Mitrović, Cvetana Krstev, and Duško Vitas. Hybrid Sentiment Analysis Framework for a Morphologically Rich Language. *Journal of Intelligent Information Systems*, 46(3):599–620, 2016b. doi: 10.1007/s10844-015-0372-5. URL <http://dx.doi.org/10.1007/s10844-015-0372-5>.
- Cristina Nicolae, Gabriel Nicolae, and Sanda Harabagiu. UTD-HLT-CG: Semantic Architecture for Metonymy Resolution and Classification of Nominal Relations. In Eneko Agirre, Lluís Màrquez, and Richard Wicentowski, editors, *SemEval '07: Proceedings of the 4th International Workshop on Semantic Evaluations*, pages 454–459, Stroudsburg, PA, USA, 2007. Association for Computational Linguistics.
- Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs Up?: Sentiment Classification Using Machine Learning Techniques. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing*, volume 10 of *EMNLP '02*, pages 79–86, Stroudsburg, PA, USA, 2002. Association for Computational Linguistics. doi: 10.3115/1118693.1118704. URL <http://dx.doi.org/10.3115/1118693.1118704>.
- Massimo Poesio and Ron Artstein. Anaphoric Annotation in the ARRAU Corpus. In Calzolari et al. (2010), pages 1170–1174. ISBN 2-9517408-6-7. URL <http://www.lrec-conf.org/proceedings/lrec2010/index.html>.
- Vassiliki Rentoumi, Stefanos Petrakis, Manfred Klenner, George A. Vouros, and Vangelis Karkaletsis. United we Stand: Improving Sentiment Analysis by Joining Machine Learning and Rule Based Methods. In Calzolari et al. (2010). ISBN 2-9517408-6-7. URL <http://www.lrec-conf.org/proceedings/lrec2010/summaries/41.html>.
- Antonio Reyes and Paolo Rosso. Building Corpora for Figurative Language Processing: The Case of Irony Detection. In *Proceedings of the 4th International Workshop on Corpora for Research on Emotion Sentiment & Social Signals (in conjunction with LREC 2012)*, pages 94–98, 2012a.
- Antonio Reyes and Paolo Rosso. Making Objective Decisions from Subjective Data: Detecting Irony in Customer Reviews. *Decision Support Systems*, 53(4):754–760, 2012b. doi: 10.1016/j.dss.2012.05.027. URL <http://dx.doi.org/10.1016/j.dss.2012.05.027>.
- Ekaterina Shutova, Simone Teufel, and Anna Korhonen. Statistical Metaphor Processing. *Computational Linguistics*, 39(2):301–353, 2013. ISSN 0891-2017. doi: 10.1162/COLI_a_00124.
- Oren Tsur, Dmitry Davidov, and Ari Rappoport. ICWSM - A great catchy name: Semi-supervised recognition of sarcastic sentences in online product reviews. In William W. Cohen and Samuel Gosling, editors, *Proceedings of the*

- Fourth International Conference on Weblogs and Social Media, ICWSM 2010, Washington, DC, USA, May 23-26, 2010.* The AAAI Press, 2010. URL <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1495>.
- Miloš Utvić. *Izgradnja referentnog korpusa savremenog srpskog jezika*. Doktorska disertacija, Univerzitet u Beogradu, Filološki fakultet, Beograd, 2014.
- Tony Veale. Detecting and Generating Ironic Comparisons: An Application of Creative Information Retrieval. In *Artificial Intelligence of Humor, Papers from the 2012 AAAI Fall Symposium, Arlington, Virginia, USA, November 2-4, 2012*, volume FS-12-02 of *AAAI Technical Report*. AAAI, 2012. URL <http://www.aaai.org/ocs/index.php/FSS/FSS12/paper/view/5557>.
- Tony Veale and Yanfen Hao. Support Structures for Linguistic Creativity: A Computational Analysis of Creative Irony in Similes. In *Proceedings of CogSci 2009, the 31st Annual Meeting of the Cognitive Science Society*, pages 1376–1381, 2009.
- Lowri Williams, Christian Bannister, Michael Arribas-Ayllon, Alun D. Preece, and Irena Spasić. The Role of Idioms in Sentiment Analysis. *Expert Systems with Applications*, 42(21):7375–7385, 2015. doi: 10.1016/j.eswa.2015.05.039. URL <http://dx.doi.org/10.1016/j.eswa.2015.05.039>.

А Прилог

Колекција текстова од 10 дигитализованих дела разноврсног жанра (дечје песме, романи, бајке, комедије и студије) којом је тестиран метод препознавања реторичке фигуре компарација.

1. Antić, Miroslav. (1965). Plavi čuperak, преузето 25.02.2016, <http://www.kotrazadragacevo.org/Knjige.html>
2. Bajke Za Decu, преузето 25.02.2016, <http://www.zadecu.com/Bajke>
3. Čopić, Branko. Orlovi rano lete, преузето 25.02.2016, <http://www.kotrazadragacevo.org/Knjige.html>
4. Danojlić, Milovan. (1966). Kako spavaju tramvaji, преузето 25.02.2016, http://borovo1234.weebly.com/uploads/1/2/6/2/12626321/milovan_danojlic_-_kako_spavaju_tramvaji.pdf
5. Martin, George R. R. Vitez sedam kraljevstava. Beograd: Laguna, 2014.
6. Nušić, Branislav. (1935). Ožalošćena porodica, преузето 25.02.2016, <https://www.scribd.com/doc/60214229/Branislav-Nusic-Ozaloscena-Porodica>
7. Orwell, Gorge. 1984. Beograd: Laguna, 2014.
8. Petrović, Mihailo. Metafore i alegorije. Beograd: Srpska književna zadruga, 1967.
9. Ršumović, Ljubivoje. Ma šta mi reče. Beograd: Laguna, 2013.
10. Stančević, Sreten. „Dečiji osmeh“. Klub pesnika “Orfej”, преузето 25.02.2016, <http://www.orfejsu.com/deciji-osmeh-sreten-stancevic-zbirka-decijih-pesama/>

Лексичка анализа двочланих термиолошких синтагми електродистрибутивног система

УДК 811.163.41'373.46

САЖЕТАК: У овом раду приказана је анализа реченичних чланова двочланих термиолошких синтагми из домена електродистрибутивног система, представљене су специфичности творбеног аспекта и размотрен је међуоднос стручне терминологије и опште лексике. Кроз анализу парадигматске равни, указано је и на одређене лексичко-семантичке законитости које у електроенергетском термиолошком систему владају на српском језику као и на промену односно постојаност њиховог реченичног облика приликом превођења на енглески језик. Направљен је посебан осврт на међујезичке утицаје, односно продор терминологије страног порекла у српски језик. Резултати добијени овим истраживањем указују на то да анализирани термини употребљени у свом синтагматском облику тешко могу наћи примену ван термиолошког система којем припадају.

КЉУЧНЕ РЕЧИ: двочлане термиолошке синтагме, електроенергетски систем, електродистрибутивни систем, порекло лексема, преводни еквиваленти.

РАД ПРИМЉЕН: 17. март 2016.

РАД ПРИХВАЋЕН: 04. мај 2016.

Тања Ивановић

tanja.ivanovic09@gmail.com

Влада Републике Србије

Канцеларија за европске интеграције

1. Увод

Као област која је у последњој деценији доживела велики успон на глобалном нивоу, енергетика је и као научна дисциплина преузела водећу улогу у друштву па је и број научних радова и речника који се објављују из ове области велики. Неки од запаженијих речника из области енергетике доступни су путем интернета. Такви су, на пример: EDF energy – Glossary of energy terms,¹ Victor Bishop-International Glossary of Biochemistry, Construction, Energy & Power

¹ <https://www.edfenergy.com/large-business/glossary>

Engineering,² Glossary of U.S. Energy Information Administration,³ Glossary of electric industry terms.⁴ Развој енергетике је, последично, праћен и растом термиолошког фонда. Из тог разлога, попис термина а касније и њихова анализа, поред општег значаја у пословању, едукацији и комуникацији, има и врло значајну улогу у смислу сагледавања утицаја страних термина (а тако и културе) на лексику српског језика и неодвојива је компонента енергетике као научне области.

Мада постоји велики број објављених истраживања која се баве, како општим принципима у раду на терминологији тако и подробнијом лингвистичком анализом термина електроенергетске струке, као подстицај на ово истраживање послужила је „Лингвистичка анализа научног стила руског и српског језика у области електротехнике“, докторска дисертација Надежде Лаиновић-Стојановић (Лаиновић-Стојановић, 1996).

С обзиром на то да су електродистрибутивни појмови, иако у врло фреквентној употреби, недовољно лексички испитани и образложени, предмет истраживања овог рада представљају двочлани електроенергетски термини на српском језику, са освртом на њихове еквиваленте на енглеском језику, у циљу да се овај лексички систем, макар и делимично, с лингвистичког аспекта приближи његовим стручним круговима.

Рад је настао као потреба привредних друштава за дистрибуцију електричне енергије Електропривреде Србије да се подзаконски акт којим се уређују рад, планирање и одржавање дистрибутивног система, а чији је званични назив Правила о раду дистрибутивног система,⁵ преведе са српског на енглески језик, поштујући правила стандардизоване стручне терминологије на оба језика.

С обзиром на то да је број радова у којима се пописују термини из области електроенергетике на српском језику врло мали,⁶ тако је и попис ових специјализованих реченичних конструкција један од циљева овог рада.

² <http://gen.lib.rus.ec/book/index.php?md5=FD402172420BD18B069D99A7DDB2FBF2>

³ <http://www.eia.gov/tools/glossary/index.cfm?id=M>

⁴ <https://www.xcelenergy.com/staticfiles/xs/Regulatory/EEIGlossaryIRPEEI2005Definitions.pdf>

⁵ Документ Правила о раду дистрибутивног система доступан је на адреси <http://www.propisonline.com/Ind0k/Legislation/52592>

⁶ Уколико изузмемо обавезу Канцеларије за Европске интеграције Владе РС да преведе комплетан корпус правних тековина ЕУу циљу придруживања Србије Европској унији, поштујући начела стручне, језичке и правнотехничке редактуре, ниједан други озбиљнији покушај пописивања термина из области енергетике није забележен. Једини битнији покушај формирања терминсистема у оквиру Електропривреде Србије своди се на публикацију коју је 1997. године издао Електроенергетски координациони центар под именом Статистичка терминологија коришћена у електропривреди на српском, енглеском и француском језику чији циљ

2. Анализа грађе

Као део лексики који обухвата скуп специјалних речи и израза из одређене научне, стручне, професионалне области, терминологија има велики значај за тачну и ефективну комуникацију у оквиру одговарајућег специјализованог језика а уједначена и прецизна терминологија пресудна је за добар превод стручних текстова (Влада Републике Србије, Канцеларија за европске интеграције, 2013). Као научна дисциплина, терминологија покрива три основна концепта, а то су: основна начела и појмови на којима се заснива проучавање термина, смернице које се користе у раду на терминологији као и скуп термина одређеног предметног поља (Castellví, 1999).

Од укупног броја који премашује хиљаду двочланих синтагми у Правилима о раду дистрибутивног система, одабрано је нешто више од сто најчешће коришћених које су приказане у прилогу на крају рада.

Анализом грађе, утврдили смо да се у српском језику најчешће јављају терминолошке синтагме типа: придев + именица. Неретко се јављају и синтагме типа: именица + именица у генитиву, али о томе ће бити речи касније.

За почетак, издвајамо двочлане синтагме типа: придев + именица, од којих су оба реченична члана страног порекла. Наводимо неколико примера:

<i>електроенергетски систем</i>	'power system'
<i>енергетска анализа</i>	'power analysis'
<i>стационарни режим</i>	'stationary regime'
<i>енергетски трансформатор</i>	'power transformer'

Као дисциплина која се бави прикупљањем, обрадом, описивањем и представљањем термина који припадају одређеном пољу примене, терминологија је задужена за проналажење одговарајућих еквивалената за стручне термине из изворног језика односно успостављање везе између термина и концепата који они представљају (Bowker, 2015). Међу анализираним лексемама значајан је број међународних ужестручних термина који су задржали своје значење из језика из ког су преузети. Највећи број лексема из анализираног корпуса води порекло из латинског језика а значајан број потиче из грчког. У малом броју случајева, лексеме које испитујемо а које потичу из латинског односно грчког језика, у српски језик су ушле посредством

је био да се створи јединствена терминологија намењена стручњацима из области статистике у Електропривреди. Самим тим, овај приручник не обухвата комплетан електроенергетски терминосистем али ипак представља позитиван покушај да се на једном месту пронађу неки од најфреквентније коришћених термина.

француског, немачког и енглеског језика. Табелом 1 приказане су лексеме страног порекла у односу на њихов укупан број у анализираној грађи, разврстане према језицима из којих воде порекло.

језик	лексеме	број речи	проц.
латински	activus, apparatus, armarium transformator, coefficient, consumere, polus, obiectum, condensator, minimalis, maximalis, differentialis, contactus, generator, factor, stationarius, combinare, frequens, directus, magnetizare, distributivus, frequentia, primarius, secundarius	24	11,0%
грчки	analysis, charakteristikos, elektron, energiea, energetikos, kriterion, metron, parallelos, pausis, phasis, synchronos, systema, thermos	13	6,0%
француски	batterie, groupe (итал. gruppo), reserve, regime (лат. regimen), electrique (грч. electron)	5	2,0%
немачки	schema (грч. schema)	1	0,5%
енглески	impedance (лат. impedire), flick	2	1,0%
укупно		45	21%

Табела 1: Лексеме страног порекла.

Поред табеларно приказаних примера лексема које имају интернационално порекло и које су се као такве одомаћиле у српском језику струке, задржавајући своје значење из језика из ког су преузете, важно је истаћи и примере у којима се јављају различита значења наизглед истих термина која, управо због своје „неусклађености“ могу довести до праве забуне приликом превођења. Такви термини називају се „лажним пријатељима“. Лажни пријатељи су речи које исто звуче али имају различита значења у два језика. Тако, термин регулатор којим се на српском језику означава уређај који служи за аутоматску регулацију напона агрегата (*напонски регулатор* – ‘voltage controller’) или брзину окретања турбине (*турбински регулатор* – ‘turbine controller’), на енглеском језику означен је лексемом ‘controller’. Насупрот томе, лексема ‘regulator’ на енглеском језику означава регулаторно тело (институцију), као што је Агенција за енергетику. Ова именица је латинског порекла а означава „уређај за аутоматско одржавање равномерног рада“ Анић и др. (2002). Сличан је случај и са термином ‘capacity’ који у електроенергетици на српском језику представља *инсталисану снагу* (а не капацитет), односно појам који се везује за електрични елемент кондензатор, као и са термином *трансформаторска станица* који свој еквивалент на енглеском језику налази у појму ‘substation’ а који се, преведен као подстанција, у српском језику везује за топлотну енергију односно грејање а не за електроенергетику. Још један пример који се често употребљава у језику струке је синтагма *управљање системом* (‘system control’), где ‘control’

не представља контролу у смислу надзора или провере већ техничко управљање системом, било да се управљање спроводи различитим управљачким уређајима (регулаторима) било да је у питању људски фактор који изводи манипулације (укључења и искључења елемената система) у систему. У том смислу, уочавамо да не постоји једнозначно пресликавање из једног језика у други као и да је терминологија страног порекла у српски стручни језик уметана уз претходну подробну анализу стручњака из дате области а ослањајући се на праксу дужу од једног века на нашим просторима. Таква терминологија је, у анализираном корпусу електроенергетске струке, данас потпуно адаптирана и одомаћена.

Највећи број терминолошких синтагми на српском језику, у корпусу којим се ово истраживање бави, има модел синтагме придев + именица, а овај модел се на енглески језик најчешће преводи моделом синтагме именица + именица, који је и најзаступљенији у енглеском језику. Тако, из наведених примера видимо да се само последња од горенаведених синтагми *стационарни режим* 'stationary regime', на енглески језик преводи истим моделом: придев + именица, док су остале преведене по моделу: именица + именица.

Уочавамо, такође, да су чести и случајеви у којима је придев страног док је именица словенског порекла:

<i>ре/активна снага</i>	're/active power'
<i>трансформаторско поље</i>	'transformer bay'
<i>ин/директно мерење</i>	'in/direct (electricity) metering'
<i>диференцијална заштита</i>	'differential protection'
<i>електрични лук</i>	'electrical arc'

Исто тако, постоје и примери синтагми у којима је придев домаћег док је именица страног порекла, мада у анализираном материјалу то није чест случај, па тако наводимо:

<i>мерна група</i>	'metering group'
<i>преносни систем</i>	'transmission system'
<i>кружна фреквенција</i>	'angular frequency'

Издајамо, и неколико примера у којима се терминолошке конструкције на српском језику састоје само од лексема домаћег порекла:

<i>пад напона</i>	'voltage drop'
<i>вршина снага</i>	'peak power'
<i>спојно поље</i>	'busbar coupler'
<i>уклопна шема</i>	'topology diagram'
<i>кратак спој</i>	'short circuit'

Анализирајући издвојене синтагме утврдили смо да се у врло малом броју случајева појављују синоними (што се може приписати прецизности и егзактности науке коју анализирамо), али, ипак, међу примерима који се јављају постоји паралелна и равноправна употреба домаћих и страних термина, па тако *називни*, *назначени* или, пак, *номинални* напон на српском језику свој еквивалент на енглеском језику налази у синтагми ‘rated voltage’. Слично је и са термином *мала* односно *мини* хидроелектрана који се на енглески језик преводи конструкцијом ‘small power plant’ или пак, примером у ком се *једнофазни* односно *монофазни* прикључак на енглески језик преводи конструкцијом ‘single phase connection’. Мада су термиолошке варијације и синоними карактеристични за употребни језик, спонтана, слободна и немотивисана употреба варијанти и лексичких синонима није у интересу ниједне струке, управо зато што доводи до недоследности у коришћењу њене терминологије (Schmitz and Straub, 2010). Стога, неопходно је да свака стручна област нормира сопствену терминологију чиме се елиминишу двосмислености и термиолошке варијације. Стварање термиолошке базе, као базе података која садржи информације о подручјима примене одређених концепата као и термина који их означавају (Melby, 2012), један је од кључних елемената у нормирању терминологије и основни предуслов за постизање високог квалитета у термиолошком раду.⁷

Међу термиолошким синтагмама домаћег порекла постоје и оне које се на енглески језик превode сажето, моноксемом, па тако издвајамо следеће примере:

<i>трансформаторска станица</i>	‘substation’
<i>разводно постројење</i>	‘switchyard’
<i>раскидач (струјног) кола</i>	‘recloser’
<i>инсталисана снага</i>	‘capacity’
<i>електрична енергија</i>	‘electricity’
<i>мерни (разводни) орман</i>	‘cubicle’

Мада се у српском језику јављају моноксеме која се на енглески језик превode двочланим синтагмама, у анализираном материјалу не издвајају се такви примери. Међутим, уочене су вишечлане синтагме у енглеском језику које своје еквиваленте на српском имају у двочланим синтагмама, попут:

⁷ Једна од највећих термиолошких база институција Европске уније је база InterActive Terminology for Europe – IATE (<http://iate.europa.eu/SearchByQueryLoad.do?method=load>), а један од позитивних примера је и Евротека, термиолошка база Канцеларије за европске интеграције Владе РС (<http://prevodjenje.seio.gov.rs/evroteka/index.php?jezik=engl>).

<i>далеководно поље</i>	‘overhead line bay’
<i>подземни вод</i>	‘underground power line’
<i>генераторски прекидач</i>	‘generator circuit breaker’
<i>спојни прекидач</i>	‘connection circuit breaker’
<i>мала електрана</i>	‘small power plant’
<i>високоуичински осигурач</i>	‘knife blade fuse’
<i>струје земљоспоја</i>	‘earth fault current’

Као што је на почетку овог поглавља речено, неретко се у српском језику јавља конструкција реченице именица + именица у генитиву која се на енглески језик преводи по моделу именица + именица. Издавајмо следеће примере:

<i>пад напона</i>	‘voltage drop’
<i>губитак снаге</i>	‘power loss’
<i>растављач снаге</i>	‘power disconnecter’
<i>предаја енергије</i>	‘electricity delivery’
<i>место прикључења</i>	‘connection point’
<i>одводник пренапона</i>	‘surge arrester’

Интересантно је да се следећи примери синтагме модела: придев + именица и именица + именица у генитиву у српском језику преводи на енглески језик по моделу партицип + именица:

<i>кофицијент свођења</i>	‘referring coefficient’
<i>погонска снага</i>	‘operating power’
<i>расклопни апарат</i>	‘switching device’
<i>мерна група</i>	‘metering group’
<i>грана магнетјења</i>	‘magnetizing branch’

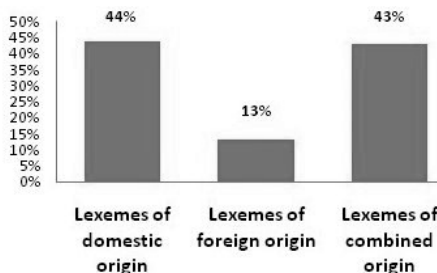
У не тако великом броју случајева али међу термилошким синтагмама које испитујемо, синтагме модела придев (трпни глаголски придев) + именица у српском језику на енглески се преводи као партицип перфекта + именица:

<i>одобрена снага</i>	‘approved power’
<i>дозвољена струја</i>	‘allowed current’
<i>комбиновани рад</i>	‘combined operation’
<i>назначена струја</i>	‘rated current’
<i>назначени напон</i>	‘rated voltage’

У односу на укупан број анализираних лексема, можемо закључити да тек петину чине термини страног порекла. И док у обрађеном корпусу 13% заузимају реченичне конструкције у ком су оба члана страног порекла, дотле се чак 44% одликује синтагмама у ком су обе лексеме домаћег порекла а

преосталих 43% се односи на синтагме комбинованог порекла (у ком је први члан страног а други домаћег порекла или обратно). У највећем броју случајева заступљена је реченична конструкција: придев + именица (чак 84%) док преосталих 16% чине конструкције: именица + именица у генитиву.

Графикон на слици 1 илуструје проценат присуства лексема домаћег и страног порекла у 107 посматраних термиолошких синтагми.



Слика 1. Приказ порекла реченичних чланова у процентима

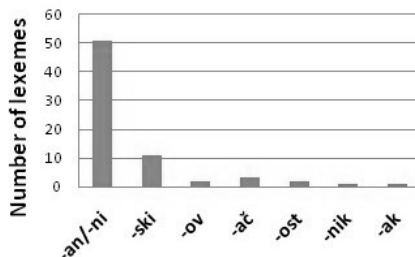
2.1 Творбена анализа

У овом одељку приказане су главне деривационе и композиционе могућности лексема у оквиру двочланих термиолошких синтагми на српском језику и представљен је инвентар афикса унутар корпуса који испитујемо.

Суфиксација Када је реч о суфиксалној анализи лексема из анализираних синтагми, у највећем броју случајева издвајају се синтагме у чијем су саставу придевски суфикси форме *-ан/-ни* и *-ски* на српском језику, док у свега два случаја налазимо и суфикс *-ов* па тако издвајамо следеће примере:

-ан/-ни: синхрони генератор, вршна снага, дистрибутивни објекат/мрежа, диференцијална заштита, дозвољена струја, електрична енергија, електрични лук, заштитни намотај, заштитни уређај, инсталисана снага, карактеристична импеданса, комбиновани рад, кружна фреквенција, конзумно подручје, контактни термометар, максимално/минимално оптерећење, мерна опрема/група итд. (укупно 51 синтагма у чијем су саставу лексеме које се формирају уз помоћ суфикса форме *-ан/-ни*);

-ски: генераторски прекидач, енергетски претварач/трансформатор, енергетска анализа, кондезаторска батерија, острвско напајање, острвски рад, погонско стање, погонска снага, трансформаторска станица/поље;



Слика 2. Суфикси

-ов: Бухолцова заштита, Тевененова импеданса.

У свега неколико примера јавља се именичка суфиксација.

-ач: растављач снаге, раскидач кола, генераторски прекидач;

-ост: сигурност напајања, критеријум сигурности;

-ник: одводник пренапона;

-ак: губитак снаге.

Графикон на слици 2 илуструје употребу суфикса у лексемама у 107 испитиваних синтагми.

Префиксација Када говоримо о префиксалној анализи лексема из анализираних синтагми на српском језику, запажамо да су доминантни префикси у овој грађи, у највећем броју случајева, домаћег порекла: *под-*, *над-*, *при-*, *пре-*, *раз-*, *без-*. Мада лексеме из издвојених синтагми представљају префиксално-суфиксалне изведенице, у примерима који следе анализирамо префиксе који их творе па тако издвајамо синтагме које се одликују префиксима и придевима домаћег порекла:

безнапонска пауза	‘voltage interruption’
надземни / подземни прикључак	‘overhead / underground connection’
пренапонска (наднапонска) / поднапонска заштита	‘overvoltage / undervoltage protection’
надпобуђени / подпобуђени режим	‘underexcitation / overexcitation regime’
привидна снага	‘apparent power’
прикључни вод	‘connection line’
разводно постројење	‘switchyard’

раставаљач снаге ‘power disconnector’

У малом броју случајева појављују се и примери у којима је префикс домаћег порекла а придев страног:

*надфреквентна/подфреквентна*⁸ ‘overfrequent / underfrequent’
заштита ‘protection’

Из обрађеног корпуса издвајамо следеће лексеме које су формиране помоћу префикса страног порекла (напомињемо да су, и префикс и придев који их творе, страног порекла), и које су, као такве, одомаћене у нашој употреби:

асинхрони генератор ‘asynchronous generator’
индиректно мерење ‘indirect electricity metering’
реактивна снага ‘reactive power’

Композиција Свака сложена реч твори се спајањем основа двеју или више посебних речи у једну реч. Из анализираних синтагми издвајамо следеће примере сложених лексема:

земљоспојна заштита ‘earth fault protection’
прекострујна заштита ‘overcurrent protection’
краткостројна заштита ‘short-circuit protection’
једнополна шема ‘single-line diagram’
једно / трофазни прикључак ‘single / three-phase connection’
високоучински осигурач ‘knife blade fuse’
полуиндиректно мерење ‘semi-indirect electricity metering’

Мада је категоријску припадност форманта *полу-* тешко одредити јер га одређени аутори сматрају префиксом, првенствено зато што често служи да делимично негира исте оне основе које префикс *не-* негира у потпуности, ограничавајући или ублажавајући значење придева, овај елемент је врло продуктиван и с њиме стално настају нове речи па смо стога овај формант и обрадили као сложеницу (Клајн, 2002, стр. 81, 116).⁹

Издавајмо и примере синтагми са сложено-суфиксалним изведеницама:

⁸ Лексеме „подфреквентан“ и „надфреквентан“, на којима није извршено једначење сугласника по звучности, преузете су из Правила о раду дистрибутивног система као званичног подзаконског акта. Стога се и у овом истраживању јавља таквом облику.

⁹ Интересантно би, пак, било питање зашто ова синтагматска конструкција није термиолошки постављена као *полудиректно мерење* уместо постојећег

далеководно поље	‘overhead line bay’
електроенергетски објекат / систем	‘power facility / system’

У овим примерима на српском језику, сложенице између првог и другог дела основе повезује спојни вокал: -о-. Једино код примера *полуиндиректно мерење* „-у- би се могло сматрати спојним вокалом додатим на основу речи *пола* или *по(л)*, уколико не прихватимо тумачење Барићеве (Barić, 1980, стр. 18–19), по којој је *полу-* префикс“ (Клајн, 2002, стр. 25).

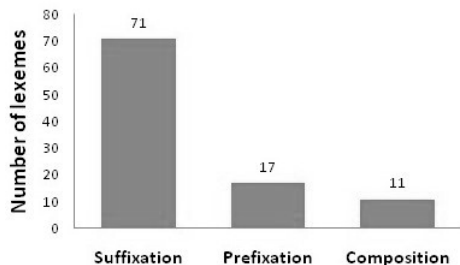
Од укупног броја анализираних лексема из обрађених синтагми лако запажамо ограничен број и репетитивни карактер како префикса тако и суфикса који их творе (имамо девет различитих облика префикса у обрађеном материјалу и седам различитих форми суфикса), за шта је у великој мери заслужна специфичност струке коју лексички испитујемо. Иако се на први поглед одликују већом разноликошћу форманата, префикси су у много мањој мери заслужни за формирање одабраног терминолошког корпуса (тек 17%), док су суфикси заступљени у чак 65% случајева. То свакако не умањује њихову важност и незаменљивост у формирању конкретне терминолошке базе. Поред тога, тек осам лексема из анализираних синтагми можемо окарактерисати чистим сложеницама које се одликују јасном мотивацијом оба дела лексеме што се, свакако, дугује егзактности посматране струке, док свега три синтагме садрже лексему која је сложено-суфиксална изведеница.

Број лексема у 107 анализираних синтагми које се творе помоћу суфикса је 71, лексема које се творе помоћу префикса је 17 а сложених лексема има 11, као што је илустровано на графикону на слици 3.

3. Лексичко-семантичке особине термина

Иако се стручни термини које испитујемо у највећем броју случајева одликују моносемантичношћу, интересантно је проучавање њихових парадигматских односа (код оних термина који имају одлике вишезначности) па се и терминолошка анализа са лексичког и семантичког становишта

полуиндиректног јер се у језичком смислу тешко уочава разлика међу њима. Међутим, та разлика ипак постоји због тога што се индиректна мерења обављају преко одређених претварача, напонских и струјних мерних трансформатора од којих, изостављањем једног, добијамо тзв. *полуиндиректно мерење*. Насупрот томе, као што и сам назив каже, директно мерење подразумева мерење без напонских и струјних мерних трансформатора преко бројила електричне енергије (у том смислу, *полудиректно мерење* се као израз у електроенергетици не користи мада би у суштини означавало исто).



Слика 3. Начини творбе речи

бави једним од најважнијих лингвистичких питања: питањима полисемије, антонимије и синонимије, којима је и посвећен овај одељак.

3.1 Полисемантичност термина

Иако је једнозначност императив код терминолошког нормирања, њено постизање је јако тешко у потпуности реализовати. С обзиром на чињеницу да су стручни термини уједно и јединице општег лексичког фонда, имајући у виду специфичност струке коју испитујемо, поставља се питање да ли ови термини, тј. терминолошке синтагме, могу, уопште, функционисати независно од терминолошког система којем припадају?

Постоје ситуације у којима, у речницима, налазимо термин који функционише искључиво у свом затвореном терминолошком систему. Уз такве речи стоји податак ком терминолошком систему припада. На пример, у Речнику српскога језика Матице српске (Матица српска, 2011) за лексему *генератор* испред њене дефиниције дат је квалификатор *физ.* чиме је потврђено да је *генератор* термин и да је јединица терминолошког система физике. У том смислу, врло тешко ћемо овакав термин искористити ван његовог терминолошког поља. Међутим, с обзиром на чињеницу да термини представљају и јединице општег језичког система у коме функционишу, према (Гортан-Премк, 2004, стр. 122–123), и са својим дефинисаним појмовним садржајем и са другим, новим који развијају у контакту са другим лексичким јединицама, када изађу из својих терминолошких система, из језика науке, када уђу у језик публицистике, и готово редовно преко њега у општи језик, почињу да имплицирају свој експлицитно дат појмовни садржај, почињу да га богате, да развијају у њему и релевантне елементе реализације, и када то, или део тога развију, почињу и да граде око себе различите лексичке системе: почињу

да развијају полисемију, чак и ону метонимијску као нпр. *генератор бројева* у лотуу.

Термини имају ограничену способност развијања вишезначности а та особина, према запажањима (Гортан-Премк, 2004, стр. 119), лежи у двојакој природи термина као знака где термин припада и општем лексичком фонду, па тако подлеже свим лексичким законитостима, између осталог и полисемантичности, али истовремено термин припада и посебном, термилошком систему па тако подлеже и његовим законитостима и тежњи ка моносемантичности, ка симетричном односу знака и појма који му је у семантичком садржају. Баш зато, термини не могу ширити своју семантичку структуру метафоричким путем а врло ретко метонимијом и синегдохом могу давати секундарна значења. Када је област коју испитујемо у питању, имамо, пак, ситуацију када се један исти термин јавља у различитим термилошким системима и са различитим појмовним вредностима. Међутим, такву појаву ипак треба разликовати од хомонимије која се, према запажањима (Шипка, 2006, стр. 61), заснива на подударности форме и неповезаности садржаја. Односно, према дефиницији (Гортан-Премк, 2004, стр. 150), хомонимија представља постојање и могућност постојања двају или више диверзних семантичких садржаја у формално, по спољашњем изгледу истоме облику.

Говорећи о узроцима настанка полисемије (Драгићевић, 2010, стр. 130) наводи да је основни разлог због којег долази до полисемије постојање знатно више предмета и појава него лексема за њих па тако, када би свака ствар око нас имала засебан назив, не бисмо могли да се споразумевамо, јер не бисмо могли да запамтимо све лексеме. Међутим, према овој ауторки, недостатак лексике не треба тумачити сиромаштвом једног језика већ ограниченим когнитивним способностима корисника језика. Деривацијом значења од основног настаје скуп значења једне лексеме који представља њену полисемантичку структуру а састоји се из примарног/основног/номинационог значења и секундарних значења (Драгићевић, 2010, стр. 131).

Од укупно 107 синтагми у овој грађи, њих 37 садржи именицу која је вишезначна и која се употребљава и ван свог термилошког система.

То су следеће именице: *апарат*, *батерија*,¹⁰ *вод*,¹¹ *генератор*, *губитак*, *грана*

¹⁰ Именица *батерија* мора се посматрати искључиво у свом синтагматском облику у ком се у анализираном корпусу јавља а то је *кондензаторска батерија*, јер само тако недвосмислено упућује на елемент у електроенергетском систему који не акумулира нити претвара енергију из једног вида у други (што је својство батерије која хемијску енергију претвара у електричну и обратно), већ има улогу у компензацији реактивне снаге.

¹¹ Означавала се *вод* којим се корисник, било електрана било потрошач, прикључује на енергетски систем, па тако ова именица остварује своје секундарно значење.

(магнећења),¹² група, енергија, заштита, коло,¹³ лук,¹⁴ место, мерење, мрежа,¹⁵ напајање,¹⁶ напон,¹⁷ објект, опрема, оптерећење, осигурач, погон, подручје, поље, рад, режим,¹⁸ свођење, сигурност, систем, снага, спој,¹⁹ станица, стање, струја, уређај, управљање, фреквенција, шема.

Ових 37 полисемантичних именица заузимају 17% од укупног броја лексема које се јављају у обрађеном корпусу. Међутим, не смемо заборавити да се велики број полисемантичних именица јавља више пута у различитим контекстима, па тако имамо више врста *заштите* (чак 11), *снаге* (такође 11), *оптерећења* (4), *струја* (5), *водова* (3), *мерења* (3), *поља* (3), *рада* (3), *режима* (3), *система* (3), *генератора* (2), *енергије* (2), *места* (2), *напајања* (2), *напона* (2), *објекта* (2), *погона* (2), *сигурности* (2), *стања* (2), *уређаја* (2) и *шема* (2).

значење које остварују према РСЈ	број лексема
примарно	16
секундарно	21
укупно	37

Табела 2: Број лексема према свом значењу у РСЈ

¹² Грана *магнећења*, у електроенергетици позната и као попречна или оточна грана. Мада се не налази у РСЈ, *магнећење* подразумева процес у ком се електрична енергија преноси магнетним доменима феромагнетног језгра електричне машине односно подразумева намагнетисавање језгра (да би електрична машина радила потребно је намагнетисати језгро).

¹³ Подразумева затворену проводну путању преко које се затвара струјни круг: струјно коло.

¹⁴ Подразумева ток струје кроз слабо проводну средину при чему се ослобађа велика количина топлоте што има јак светлосни ефекат.

¹⁵ Према РСЈ, значење којим се ми бавимо у овој анализи налази се на 5. месту, што значи да је лексема из опште употребе проширила своју семантичку структуру метафоричким путем и добила терминолошко значење

¹⁶ Нема га као именице у РСЈ. Од гл. *напајати*, које у свом примарном облику представља несвршени глагол према напојити се, а у свом секундарном значењу подразумева снабдевати се, пунити се (водом, струјом и сл.) према РСЈ.

¹⁷ Разлика у потенцијалу између двеју тачака у електричном колу или пољу, која омогућава протицање струје кроз затворен круг.

¹⁸ Односи се на одређено радно стање у систему. У анализираном корпусу имамо *стационарни режим*, *подпобуђени режим*, *надпобуђени режим*.

¹⁹ У обрађеном материјалу јавља се и именица које нема у РСЈ: *земљоспој* (*струје земљоспоја*), а која подразумева вредност струје која се јавља када један од три фазна проводника додирне нулти потенцијал.

Иако лексеме из реченичних синтагми могу коегзистирати у општем лексичком систему када се издвоје из синтагми које испитујемо, ту способност губе када су употребљене у синтагматском односу језика струке јер тада недвосмислено указују на одређену појаву унутар електроенергетског система, место или целину уређаја и инструмената неопходних за функционисање система. Самим тим можемо закључити да је проценат затворености анализираног терминолошког система висок. Табелом 2 приказан је број полисемантичних лексема према значењу које остварују у Речнику српског језика Матица српска (2011).

3.2 Антонимски однос лексема

И док се, у синтагмама које испитујемо, полисемија јавља искључиво на именицама, дотле се антонимија јавља на придевима, па тако издвајамо:

<i>минимално оптерећење</i>	‘minimum load’ –
<i>максимално оптерећење</i>	‘maximum (peak) load’
<i>подземни вод/прикључак</i>	‘underground power line/connection’ –
<i>надземни вод/прикључак</i>	‘overhead power line/connection’
<i>подпобуђени режим</i>	‘underexcitation regime’ –
<i>надпобуђени режим</i>	‘overexcitation regime’
<i>подфреквентна заштита</i>	‘underfrequent protection’ –
<i>надфреквентна заштита</i>	‘overfrequent protection’
<i>пренапонска (или: наднапонска)</i>	‘overvoltage protection’ –
<i>заштита –</i>	
<i>поднапонска заштита</i>	‘undervoltage protection’ ²⁰

Издвојени примери лексема, било да представљају степеневане било нестепеневане антониме, у овом односу припадају истој лексичко-семантичкој групи односно имају исто основно значење речи. У већини случајева, издвојени придеви су сложене структуре, настали творбом речи, што је случај и са њиховим антонимима.

Поред тога, интересантни су и примери у којима синтагме представљају лексичке антониме али имајући у виду различит начин функционисања у електроенергетском систему, они немају супротстављену функцију, па их можемо окарактерисати као комплементарне антониме,²¹ попут:

²⁰ Ради се о две врсте напонске заштите па сходно томе, у антонимски однос не може се ставити: напонска – поднапонска, или напонска – пренапонска заштита.

²¹ Одражавајући, у основи, сличност у различитостима дела ванјезичке стварности, а потом и дела лексикона који је одсликава, антонимија оличава супротност смисла између две лексеме. У складу с природом те супротности, могуће је разликовати

<i>једнофазни прикључак</i>	‘single phase connection’ –
<i>трофазни прикључак</i>	‘three-phase connection’
<i>примарна струја</i>	‘primary current’ –
<i>секундарна струја</i>	‘secondary current’
<i>примарни намотај</i>	‘primary winding’ –
<i>секундарни намотај</i>	‘secondary winding’)
<i>директно мерење</i>	‘direct electricity metering’ –
<i>индиректно мерење</i>	‘indirect electricity metering’ –
<i>полуиндиректно мерење</i>	‘semi-indirect electricity metering’
<i>активна снага</i>	‘active power’ –
<i>реактивна снага</i>	‘reactive power’
<i>синхрони генератор</i>	‘synchronous generator’ –
<i>асинхрони генератор</i>	‘asynchronous generator’

У свом прототипском испољавању, комплементарна антонимија претпоставља немогућност поређења посматраних стања ствари (Prčić, 1997, стр. 105).

Међу анализираним синтагмама, запажамо пример: *дозвољене струје* који заправо означава *максимално дозвољене струје* ‘maximum allowed current’ у коме се *максимално* подразумева па се често и избацује из синтагме, чак и у писаним текстовима, и која нема свој антоним, односно не постоје *недозвољене струје* или *минимално дозвољене струје*. Исти је случај и са примером *одобрена снага* чији би антоним у општој лексици био придев: *неодобрена*, док он, као такав, у области коју испитујемо не постоји. У анализираној грађи јавља се и пример: *прекострујне заштите*, док, аналогно претходним примерима, *исподструјни* не постоји а исти је случај и са *високоучинским осигурачем* јер *нискоучински* не постоји.

Антонимија се индукује у основним, једнореферентним лексемама, управо у оним лексемама које индукују полисемију и деривацију и које искључују синонимију (као секундарну лексичку појаву). Ово упућује на закључак да је и антонимија, заједно са варирањем (полисемантичким и деривационим), основни механизам у организацији лексичког система (Гортан-Премк, 2004, стр. 149). А као што закључујемо, свакако, и електроенергетског.

У односу на укупан број примера у корпусу од 107 двочланих термиолошких синтагми које испитујемо у овом раду, можемо издвојити 12 антонимских парова (укупно 25 термиолошких синтагми), што значи да се антонимија јавља у 23% случајева у односу на укупан број синтагми. То значи да је антонимија у приличној мери заслужна за формирање овог терминосистема. Примери синонимије у електродистрибутивној, односно електроенергетској,

најмање пет основних врста антонимије – скаларну, комплементарну, дијаметралну, реципрочну и реверзивну (Prčić, 1997, стр. 102).

терминологији су ретки, можда управо због специфичности дисциплине која тежи што већој прецизности и одређености, а у обрађеном материјалу можемо издвојити два, у ком су обе лексеме домаћег порекла:

пренапонска заштита – ‘overvoltage protection’
наднапонска заштита
грана магнетезња – ‘magnetizing branch’
попечна грана – оточна грана

За преостале три: ‘rated voltage’, ‘small power plant’, ‘single phase connection’, за које смо већ утврдили да имају своје синониме, паралелно се користе и домаћи и страни термини у преводу, о чему је већ било речи.

Од 107 примера анализираних синтагми, табелом 4 пружен је приказ испитиваних лексема према творбеним способностима.

типови термилошке лексике према начину творбе			број	проц.
лексема				
сложене			8	7.5%
сложено-суфиксалне изведенице			3	3.0%
префиксалне (префиксално-суфиксалне) творенице: над-, под-, раз-, при-, ин-, без-, а-, ин-, ре-			18	17.0%
суфиксалне творенице	именички суфикси		7	6.5%
	придевски суфикси:	-an/-ni	49	46.0%
		-ski	11	10.0%
		-ov	2	1.9%
полисемантичне лексеме			37	34.5%
антоними			25	23.0%
синоними			5	4.7%

Табела 3. Термилошке синтагме приказане према броју лексема и њиховом проценту из одабраног корпуса.

4. Међујезички утицаји односно продор терминологије из других језика

Мада савремени стручњаци електроенергетике сматрају да за сваки стручни термин из страног језика треба пронаћи одговарајући еквивалент на српском језику, ипак постоји одређени број термина који су међународног карактера и као такви, већ су инкорпорирани у лексику савременог српског језика и прилагођени његовом писму и структури. То, свакако, има небројене предности јер се таквом унификацијом термина олакшава комуникација међу научницима

и стручњацима на различитим предавањима, конференцијама или, пак, при пословним контактима а никако не смемо заборавити ни стручну литературу чије је коришћење интернационализацијом термина у великој мери олакшано. Ако се узме у обзир и чињеница да је језик живи процес, онда је оваква појава сасвим природна и очекивана, поготово за доба у којем живимо а које се, између осталог, одликује и глобалном повезаношћу удаљених народа и језика. Поред тога, доминација утицаја језика народа чија наука у одређеном тренутку предњачи један је од кључних разлога који узрокује преузимање термина.

Када је српски језик у питању, приметан је јак утицај енглеског језика у последњих неколико декада, мада је уочљиво и присуство одређених термина који су, слободно речено, некритички усвојени. Охрабрујући су, пак, покушаји савремених стручњака електроенергетике да се стручни термини на енглеском језику замене домаћим терминима који би били у духу српског језика. Један од истакнутијих примера некритички усвојених термина је термин ‘recloser’ који се на српски језик најпре преводио као *интелигентни линијски прекидач* а потом и као *уклопник* (мада је често задржавао и свој изворни облик). С обзиром на чињеницу да енергетика као стручна област обилује уређајима који прекидају електрична кола а која се називају прекидачима јасно је да *интелигентни линијски прекидач* често није недвосмислено указивао на предмет који је представљао па је стога било неопходно направити разлику у односу на остале. Тако је и прихваћен израз *раскидач (струјног) кола* као његов еквивалент на српском језику.

5. Закључак

Мада се свака научна област може похвалити засебном терминологијом, није велики број оних које су доживеле толику експанзију последњих година, какав је случај са енергетиком. Упоредо са њеним развојем, а прилагођавајући се потребама савременог друштва и кретања, јавља се и нова, специфична лексика којом се, посредством различитих медија, истовремено обогаћује и наш стандардни језик.

С обзиром на специфичност ове научно-техничке гране коју карактерише универзална тежња ка што прецизнијем и егзактнијем, ту особину налазимо и на морфолошком плану на ком се свака синтагма, употребљена у јединици (или множини) на исти начин тј. истим бројем преводи на енглески језик. Одступања од опште прецизности налазимо у примерима двочланих синтагми на српском језику које се преводе моноклексомом на енглески или у примерима вишечланих конструкција у енглеском које своје еквиваленте налазе у двочланим синтагмама на српском језику.

Из анализе разматраних примера можемо закључити да делови синтагми (стандардне лексеме) могу коегзистирати у општем лексичком систему када се издвоје из синтагми које испитујемо, самостално, али да се таква њихова способност губи када су употребљени у синтагматском односу језика струке, јер као такве нису својствене општем говору и недвосмислено указују на одређену појаву унутар електроенергетског система, место или целину уређаја неопходних за функционисање система. У прилог томе иде и чињеница да се такве речи употребљавају уз именицу и служе као атрибути па тако образују именичке синтагме карактеристичне за електроенергетску струку. Из тог разлога, врло је тешко ове термилошке синтагме употребити независно од термилошког поља којем припадају па стога оне и остају затворене у својим термилошким круговима. Овакве синтагме представљају јединство појма или представе, чинећи тако унутрашњу „форму“ језика (Белић, 1958), а ми слободно можемо додати, језика струке.

У односу на укупан број анализираних лексема из синтагми у овом раду, тек петину чине термини страног порекла (у највећем броју случајева преузети су директно из латинског и грчког језика, док у нешто мањем броју из француског, немачког и енглеског језика). Од укупно 107 термилошких синтагми које чине анализирани корпус, 13% заузимају реченичне конструкције у ком су оба члана страног порекла, чак 44% одликује се синтагмама у ком су обе лексеме домаћег порекла док се 43% односи на синтагме комбинованог порекла. Иако је број префикса односно суфикса сведен на тек неколико различитих форманата (имамо девет различитих облика префикса у анализираном материјалу и седам различитих облика суфикса), префикси су у много мањој мери заслужни за формирање одабраног термилошког корпуса (17%) док су суфикси заступљени у чак 65% случајева, у односу на укупан број анализираних лексема. Па ипак, мада њихова деривациона улога није једнака у формирању конкретне термилошке базе, она је подједнако важна. Поред тога, тек 11 лексема из ове грађе одликује се сложеностом структуром коју карактерише јасна мотивација оба дела лексеме. Ограничен број лексема из синтагми анализираних у овом раду (њих 37), показује и особину полисемантичности. Карактеристику затвореног терминосистема потврђују и примери синонимије из разматраног материјала који су веома ретки (тек пет лексема показало је ову особину), док се антонимија показала као један од значајнијих механизма у организацији лексичког система струке коју испитујемо (присутна је чак 23% у корпусу који анализирамо).

Како термилошке синтагме које анализирамо одражавају свеобухватно знање из области електроенергетике које својим семантичким обележјима пружају језичку и појмовну тачност, и поред тога што показују тежњу за стварањем антонимије, у малом броју случајева синонимије или што пружају отпор полисемији, ово истраживање је показало да се њихова ограниченост на

сопствени терминосистем не доводи у питање и да другачије односе не могу успоставити у стандардној, па чак ни у неформалној употреби.

С обзиром на то да је рађена на малом узорку терминолошких синтагми, ова лингвистичка анализа пружа само делимичан одговор на лексичка питања која су овим радом обрађена. Иако скроман допринос, значајно је полазиште и добар темељ за даљи преводилачки рад на званичним актима у електродистрибутивној делатности као што су техничке препоруке и интерни стандарди којима се допуњују Правила о раду дистрибутивног система. Поред тога, представља и солидан подстицај за разматрање лексичких питања у стручним круговима, односно прецизно одређивање значења одређених термина као и стандардизацију терминологије унутар специјализоване струке.

Литература

- Anić, Vladimir, Dunja Brozović-Rončević, Ivo Goldstein, Slavko Goldstein, Ljiljana Jojić, Ranko Matasović i Ivo Pranjković. *Hrvatski enciklopedijski rječnik*. Zagreb : Novi liber, 2002.
- Barić, Eugenija. *Imeničke složenice i neprefiksane i nesufiksane tvorbe*. Zagreb : Liber, 1980.
- Белић, Александар. *О језичкој природи и језичком развоју: лингвистичка испитивања*. Београд : Нолит, 1958.
- Bowker, Lynne. “Terminology and translation”. In: *Handbook of Terminology*, edited by Hendrik J. Kockaert and Frieda Steurs, Vol. 1, pp 304–323, Amsterdam: John Benjamins Publishing Company, 2015.
- Cabré, M Teresa. *Terminology: Theory, methods and applications*. Amsterdam : John Benjamins Publishing, 1999.
- Драгићевић, Рајна. *Лексикологија српског језика*. Београд : Завод за уџбенике и наставна средства, 2010.
- Гортан-Премк, Даринка. *Полисемја и организација лексичког система у српском језику*. Београд : Завод за уџбенике и наставна средства, 2004.
- Клајн, Иван. *Творба речи у савременом српском језику, први део: слагање и префиксација*. Београд : Завод за уџбенике и наставна средства, 2002.
- Лаиновић-Стојановић, Надежда. „Лингвистичка анализа научног стила руског и српског језика у области електротехнике (лексички ниво)“. Докт. дисертација, Универзитет у Београду, 1996.
- Матица српска. *Речник српског језика*. Нови Сад : Матица српска, 2011.
- Melby, Alan K. “Terminology in the age of multilingual corpora”. *The Journal of Specialised Translation*, 18(2012):7–29. http://www.jostrans.org/issue18/art_melby.php.

- Prčić, Tvrtko. *Semantika i pragmatika reči*. Sremski Karlovci; Novi Sad : Izdavačka knjižarnica Zorana Stojanovića, 1997.
- Schmitz, Klaus-Dirk and Daniela Straub. *Successful terminology management in companies: practical tips and guidelines: basic principles, implementation, cost-benefit analysis and system overview*. Stuttgart : Tekom, 2010.
- Шипка, Данко. *Основи лексикологије и сродних дисциплина*. Нови Сад : Матица српска, 2006.
- Влада Републике Србије, Канцеларија за европске интеграције. *Приручник за превођење правних аката Европске уније*. Београд: Република Србија, Влада, Канцеларија за европске интеграције, 2013.

Прилог – Терминолошке синтагме на српском и енглеском језику са пореклом речи²²

Терминолошка синтагма на српском	Порекло прве лексеме	Порекло друге лексеме	Терминолошка синтагма на енглеском
1 активна снага	лат. <i>activus</i>		active power
2 асинхрони генератор	грч. <i>synchronos</i>	лат. <i>generator</i>	asynchronous generator
3 безнапонска пауза		грч. <i>pausis</i>	voltage interruption
4 Бухолцова заштита			Buchholz protection
5 високоучински осигурач			knife blade fuse
6 вршна снага			peak power
7 вршно оптерећење			peak load
8 генераторски прекидач	лат. <i>generator</i>		generator circuit breaker
9 грана магнећења		нлат. <i>magnetizare</i>	magnetizing branch
10 губитак снаге			power loss
11 далеководно поље			overhead line bay
12 директно мерење	лат. <i>directus</i>		direct electricity metering
13 дистрибутивна мрежа	нлат. <i>distributivus</i>		distribution network
14 дистрибутивни објекат	нлат. <i>distributivus</i>	лат. <i>obiectum</i>	distribution facility
15 диференцијална заштита	нлат. <i>differentialis</i>		differential protection
16 дозвољена струја			(maximum) allowed current

²² Колоне *Порекло прве лексеме* и *Порекло друге лексеме* су попуњене само уколико лексема из терминолошке синтагме има страно порекло.

Терминолошка синтагма на српском	Порекло прве лексеме	Порекло друге лексеме	Терминолошка синтагма на енглеском
17 електрична енергија	фр. <i>electrique</i>	гр. <i>energeia</i>	electricity
18 електрични лук	фр. <i>electrique</i> од грч. <i>elektron</i> ;		electrical arc
19 електроенергетски објекат	грч. <i>elektron</i> ; грч. <i>energetikos</i>	лат. <i>obiectum</i>	power facility
20 електроенергетски систем	грч. <i>elektron</i> ; грч. <i>energetikos</i>	гр. <i>systema</i>	power system
21 електроенергетска анализа	грч. <i>energetikos</i>	грч. <i>analysis</i>	power analysis
22 енергетски претварач	грч. <i>energetikos</i>		power converter
23 енергетски трансформатор	грч. <i>energetikos</i>	нлат. <i>transformator</i>	power transformer
24 заштитини намотај			protective winding
25 заштитини уређај			protective device
26 земљоспојна заштита			earth fault protection
27 индиректно мерење	лат. <i>indirectus</i>		indirect electricity metering
28 инсталисана снага	срлат. <i>installatio</i>		capacity
29 једнополна шема	лат. <i>polus</i>	нем. <i>schema</i>	single-line diagram
30 једнофазни прикључак	грч. <i>phasis</i>		single phase connection
31 карактеристична импеданса	грч. <i>charakteristikos</i>	енгл. <i>impedance</i> од лат. <i>impedire</i>	characteristic impedance
32 коефицијент свођења	нлат. <i>coefficiens</i>		referring coefficient
33 коефицијент фликера	нлат. <i>coefficiens</i>	енгл. <i>flick</i>	flicker coefficient
34 комбиновани рад	лат. <i>combinare</i>		combined operation
35 кондензаторска батерија	нлат. <i>condensator</i>	фр. <i>batterie</i>	capacitor battery
36 конзумно подручје	лат. <i>consumere</i>		consumption area
37 контактни термометар	лат. <i>contactus</i>	грч. <i>thermos</i> грч. <i>metron</i>	contact thermometer
38 кратак спој			short circuit
39 краткоспојна заштита			short-circuit protection
40 критеријум сигурности	грч. <i>kriterion</i>		security criterion
41 кружна фреквенција		лат. <i>frequentia</i>	angular frequency
42 максимално оптерећење	нлат. <i>maximalis</i>		maximum / peak load
43 мала електрана		фр. <i>electrique</i>	small power plant
44 мерна група		фр. <i>groupe</i> од итал. <i>gruppo</i>	metering group

Терминолошка синтагма на српском	Порекло прве лексеме	Порекло друге лексеме	Терминолошка синтагма на енглеском
45 мерна опрема			metering equipment
46 мерни (разводни) орман		срлат. <i>armarium</i>	cubicle
47 мерни намотај			metering winding
48 мерни уређај			metering device
49 мерно место			metering point
50 место прикључења			connection point
51 минимално оптерећење	нлат. <i>minimalis</i>		minimum load
52 надземни вод			overhead line
53 надземни прикључак			overhead connection
54 надпобуђени режим		фр. <i>regime</i> од лат. <i>regimen</i>	overexcitation regime
55. надфреквентна заштита	лат. <i>frequens</i>		overfrequent protection
56 називни напон назначени напон номинални напон			rated voltage
57 назначена струја			rated current
58 напонска заштита			voltage protection
59 одводник пренапона			surge arrester
60 одобрена снага			approved power
61 острвски рад			island operation
62 острвско напајање			island supply
63 пад напона			voltage drop
64 паралелни рад	грч. <i>parallelos</i>		parallel operation
65 погонска снага			operating power
66 погонско стање			operating state
67 подземни вод			underground line
68 подземни прикључак			underground connection
69 поднапонска заштита			undervoltage protection
70 подпобуђени режим		фр. <i>regime</i> од лат. <i>regimen</i>	underexcitation regime
71 подфреквентна заштита	лат. <i>frequens</i>		underfrequent protection
72 полуиндиректно мерење	лат. <i>directus</i>		semi-indirect electricity metering
73 поремећени погон			operation disturbance
74 поуздан погон			reliable operation
75 предаја (електричне) енергије		грч. <i>energetikos</i>	electricity delivery

Терминолошка синтагма на српском	Порекло прве лексеми	Порекло друге лексеми	Терминолошка синтагма на енглеском
76 прекострујна заштита			overcurrent protection
77 пренапонска заштита			overvoltage protection
наднапонска заштита			
78 преносни систем		грч. <i>systema</i>	transmission system
79 привидна снага			apparent power
80 прикључни вод			connection line
81 примарна струја	лат. <i>primarius</i>		primary current
82 примарни намотај	лат. <i>primarius</i>		primary winding
83 разводно постројење			switchyard
84 раскидач (струјног) кола			recloser
85 расколни апарат		лат. <i>apparatus</i>	switching device
86 растављач снаге			power disconnector
87 реактивна снага	лат. <i>reactivus</i>		reactive power
88 резервна заштита	фр. <i>reserve</i>		reserve protection
89 секундарна струја	лат. <i>secundarius</i>		secondary current
90 секундарни намотај	лат. <i>secundarius</i>		secondary winding
91 сигурност напајања			security of supply
92 синхрони генератор	гр. <i>synchronos</i>	лат. <i>generator</i>	synchronous generator
93 снага трансформатора		нлат. <i>transformator</i>	transformer power
94 спојни прекидач			connection circuit breaker
95 спојно поље			usbar coupler
96 стационарни режим	клат. <i>stationarius</i>	фр. <i>regime</i> од лат. <i>regimen</i>	stationary regime
97 струје земљоспоја			earth fault current
98 струјно оптерећење			current load
99 Тевененова импеданса		енгл. <i>impedance</i> од лат. <i>impedire</i>	Thevenen's impedance
100 трансформаторска станица	нлат. <i>transformator</i>		substation
101 трафо (трансформаторско) поље	нлат. <i>transformator</i>		transformer bay
102 трофазни прикључак	грч. <i>phasis</i>		three-phase
103 уклопна шема		нем. <i>schema</i>	connection topology diagram
104 уклопно стање			topology condition
105 управљање системом		грч. <i>systema</i>	system control
106 фактор поремећаја	лат. <i>factor</i>		disturbance factor
107 фактор снаге	лат. <i>factor</i>		power factor

Библиометријска и цитатна анализа часописа *SPATIUM*

УДК 001.3:050 SPATIUM "2009/2015"

САЖЕТАК: Рад представља библиометријску и цитатну анализу часописа *SPATIUM* за период од 2009. до 2015. године. У задатом периоду *Spatium* је излазио два или три пута годишње, од броја 19 до броја 34 тако да је овом студијом обухваћено 16 свезака, односно 145 радова. Библиометријска анализа часописа рађена је *de visu* (са публикацијом у руци), чиме је обезбеђена аутентичност наведених података. Анализом је приказан број радова, категоризација радова, број аутора, афилијација аутора, заступљеност аутора из различитих држава, просечан број референци по раду, дужина текстова, учесталост кључних речи, укупан број рецензената и њихова међународна заступљеност. Цитираност часописа урађена је на основу података преузетих из претраживача и цитатних база Scindeks, Google scholar и Scopus. У раду су представљене разлике у анализи цитираности научног часописа *SPATIUM* на примеру једног рада.

КЉУЧНЕ РЕЧИ: научни часопис, *Spatium*, библиометријска анализа, цитатна анализа, Србија

РАД ПРИМЉЕН: 24. фебруар 2016.

РАД ПРИХВАЋЕН: 28. март 2016.

Милена Милинковић

milenam@iaus.ac.rs

Институт за архитектуру
и урбанизам Србије

1. Увод

Један од основних услова континуитета и напредовања сваке научне области у савременим условима, свакако је поуздан, квалитетан и редован научни часопис. Такав часопис правовремено информише о најновијим достигнућима у оквиру конкретне научне области.

И док је штампана публикација (прво књига, а потом и часопис) током више векова била доминантни извор научних информација, а резултати истраживања

бележени на папиру, развојем модерне технологије, од краја XX века све заступљеније су електронске публикације, како оне објављене у електронској форми као посебна физичка целина (CD, DVD), тако и публикације које су доступне на Интернету¹.

У савременом свету се сматра да научна информација није релевантна ако није објављена. Зато ће улогу универзалног комуникационог канала и медијатора у релевантном научном окружењу још дуги низ година имати научни часописи, било да су штампани или у електронском облику. Да би неки рад био публикован у научном часопису треба да добије позитивну оцену рецензента, а то подразумева да мора да задовољи унапред постављене критеријуме који се односе на оригиналност и квалитет текста. Рецензенте из одговарајуће научне области, која се поклапа са темом рада, бира уредништво часописа и то на основу научне репутације коју они имају у научним круговима.

Научни часопис представља медиј за чување научних информација, има значајну улогу у формирању научног ауторитета, представљању резултата истраживања, евалуацији стручног и научног доприноса и контроли квалитета научног рада. Радови објављени у научном часопису представљају најзначајнији канал научне комуникације. Зато улога научног чланка није само да пренесе научну информацију о готовом и непромењивом знању, већ да својим садржајем подстакне научнике на размену идеја и нека даља истраживања (Вучковић, 2009).

2. Часопис *SPATIUM*

У издању Института за архитектуру и урбанизам Србије (у даљем тексту: Институт) часопис *SPATIUM*² излази већ 19 година. У периоду од 1997. до 2015. године објављена су 34 броја, односно 31 свеска, са 261 радом и 23 прилога (приказа књига, приказа конференција, некролога, обавештења о симпозијумима, преведених раније објављених радова итд.). Од укупног броја свезака 28 је појединачних, а преостале 3 свеске су двоброји (табела 1).

¹ У Закону о обавезном примерку стоји „да је „електронска публикација” публикација објављена у електронској форми као посебна физичка целина (дискета, CD, DVD и сл.), публикација која је доступна на Интернету и публикација припремљена за штампу у формату који одговара међународним стандардима универзалне доступности информација” (Закон о обавезном примерку публикација („Службени гласник РС”, бр. 52/2011), преузето 15.02.2015, <http://bds.rs/dokumenti/Zakon%20obavezni%20primerak%202011.pdf>).

² *Spatium* / editor in chief Miodrag Vujošević. (Belgrade: Institute of Architecture and Urban & Spatial Planning of Serbia, 1997-), br. 1-34 (1997-2015)

Година	бр.	Година	бр.
1997	1, 2	1998	3, 4
1999	5	2000	6
2001	7	2002	8
2003	9	2004	10, 11
2005	12	2006	13/14
2007	15/16	2008	17/18
2009	19, 20, 21	2010	22, 23
2011	24, 25, 26	2012	27, 28
2013	29, 30	2014	31, 32
2015	33, 34		

Табела 1. Периодичност излагања

До 2009. године сви радови су имали само по једну позитивну рецензију и нису били сврставани у неку од категорија. Од 2009. године, ступањем на снагу *Акта о уређивању научних часописа*³, знатно су пооштрени критеријуми уређивања часописа. Уредништво је почело са стриктном применом *Правилника о поступку и начину вредновања и квантитативном исказивању научноистраживачких резултата истраживача*⁴ (у даљем тексту: Правилник) и од тада су критеријуми за објављивање радова позитивна мишљења два рецензента за научне радове и једна позитивна рецензија за стручне радове. Поштује се узајамна анонимност аутора и рацензената, а рецензије раде најеминентнији стручњаци из конкретне научне области како из земље тако и из иностранства.

У време када је покренут, 1997. године, часопис је замишљен као научно гласило пре свега за научне раднике Института и њихове спољне сараднике. Како је у периоду међународне изолације Србије била отежана сарадња са колегама из иностранства, покретање једног оваквог научног часописа требало је да омогући превазилажење овог проблема и успостављање лакше комуникације и научне и стручне размене мишљења.

³ Министарство за науку и технолошки развој донело је Акт о уређивању научних часописа (Ев. број: 110-00-17/2009-01, од 09.07.2009) са циљем да помогне уредништвима да побољшају квалитет и унапреде домаћу научну периодику чиме би се обезбедила већа укљученост часописа у систем научних информација како на домаћем тако и на међународном нивоу. Преузето 19.01.2016, http://kobson.nb.rs/upload/documents/MNTR/Dokumenti/akt_o_uredjivanju_casopisa.pdf

⁴ Правилник о поступку и начину вредновања, и квантитативном исказивању научноистраживачких резултата истраживача („Сл. гласник РС“, бр. 38/2008)

У часопису су заступљени радови из различитих области: просторног планирања, урбанизма, архитектуре, пејзажне архитектуре, геодезије, одрживог развоја и заштите животне средине, становања, обнове и изградње градова, комунално-стамбене проблематике, културног и природног наслеђа, инвестиционе изградње, информационих система, стратешког управљања и др.

Часопис *SPATIUM* излази на енглеском језику и налази се у категорији М24 (национални часопис међународног значаја) према категоризацији домаћих научних часописа за област Саобраћај, урбанизам и грађевинарство. Као потврда квалитета и међународног значаја часописа јесте и чињеница да је Министарство науке Републике Србије октобра 2010. године уврстило овај часопис у DOI⁵ (Digital Object Identifier) систем који омогућава идентификацију документа у електронском облику и успостављање сталне везе до Интернет странице на којој се оригинални документ налази. DOI бројеви додељени су свим радовима од 2002. године. Управо захваљујући DOI броју радови, њихови аутори и часопис у којима се публикују бивају „видљивији“ за стручну и научну јавност, што значајно доприноси међународној афирмацији аутора и самог часописа.

Од 2012. године Институт је у уговорним обавезама са *Versitom* као суиздавачем часописа *SPATIUM*. *Versita* је електронска база европских издавача научних и стручних часописа (Central European Science Publishers)⁶, која нуди онлајн технологију за приказ садржаја часописа, сажетака радова, цитираних референци, и радова у формату пуног текста, помоћ у евалуацији и глобалну промоцију часописа што би требало да резултира знатним повећањем видљивости, читаности и цитираности. Само у 2013. години *Versita* је објавила 15.000 чланака у отвореном приступу. Од почетка 2014. године нов власник *Versite* је de Gruyter и компанија мења име у *de Gruyter Open*⁷ са планом да у наредном периоду додатно прошири активност ван Европе у САД и Азију.

Крајњи циљ уредништва је улазак часописа *SPATIUM* на Thomson Reuters листу. Сваке године Thomson Reuters процењује око 2.000 нових часописа

⁵ Дигитални идентификатор објекта (DOI) је низ карактера који се користи за јединствену идентификацију електронских докумената. DOI број за неки документ је трајан и стога стабилнији метод реферирања документа него URL.

⁶ „Од свог оснивања компанија *Versita*, у власништву је г. Jaceka Ciesielskog, са седиштем у Варшави, настојала да буде водећи комерцијални издавач научних часописа Средње Европе. То је настојала да постигне, с једне стране, покретањем нових средњеевропских часописа, а са друге стране, претварањем већ реномираних часописа из земаља Средње Европе у онлине верзију уз услов да часопис излази на енглеском језику, што јој је у великој мери успело, пре свега, у Пољској, а затим и у Словачкој, Србији, Чешкој и Хрватској“ (Tóth, 2007)

⁷ <http://www.degruyter.com/view/j/spat>, преузето 02.09.2014.

за реферисње у Web of Science⁸ (WoS), а ниво пролазности је до 10–12%. Та евалуација је бесплатна, и обухвата стриктно дефинисане критеријуме. Након почетног пријављивања, потребно је послати наредне три свеске непосредно након објављивања. По пријему три узастопна броја часописа, уредници одговарајуће базе података врше процену публикације.

Радови објављени у часопису *SPATIUM* доступни су у формату пуног текста преко следећих база података, каталога и сервиса:

- doiSerbia⁹ – за радове објављене у периоду 2002–2015. године;¹⁰
- Academic Journals Database¹¹ – за радове објављене у периоду 2002–2013. године;
- Srpski citatni indeks¹² – за радове објављене у периоду 2000–2013. године;

⁸ „Web of Science®“, на који је претплаћена академска заједница у Србији садржи три основна цитатна индекса: Science Citation Index expanded SCI база података из области природних, биомедицинских и техничких наука, Social Sciences Citation Index SSCI база података која обухвата часописе из области друштвених наука; Arts & Humanities Citation Index AHCI база часописа из уметности и хуманистичких наука. Од октобра 2008. године поседује и базе података за зборнике: Conference Proceedings Citation Index Science (CPCIS) од 2001. године до данас, цитатни индекс за природне и техничке науке, Conference Proceedings Citation Index Social Science & Humanities (CPCISSH) цитатни индекс за друштвене и хуманистичке науке од 2001. године до данас Journal Citation Report (JCR) је посебна база података у оквиру платформе Web of KnowledgeSM. Настала је обрадом резултата из наведених цитатних индекса. JCR садржи по датке о часописима сврстане у тематске категорије, у оквиру којих су ранжирани по импакт фактору (енгл. ImpactFactor IF) или фактору утицаја“ (Антонић и др., 2009)

⁹ „DOI Србија је збирка научних часописа који излазе у Србији. Дигитална архива пуних текстова обухвата период од 2002. године. Часописи су обрађени од корица до корица. Све реферисане часописе издају професионалне асоцијације, а неки од њих су индексирани и у WoS и CA. Метаподаци, као и пуни текстови, урађени су према OAI-PMH, и могу бити преузети без надокнаде“, преузето 21.01.2016, <http://www.digitallibrary.eu/te14/collection/a0155?locale=sr>

¹⁰ преузето 01.02.2016, <http://www.doiserbia.nb.rs/journal.aspx?issn=1450-569X>

¹¹ Преглед садржаја, сажетак, као и радова у формату пуног текста објављених у часопису *SPATIUM* доступан је на порталу Academic Journals Database, каталогу научних публикација из различитих подручја, преузето 01.02.2016, <http://journaldatabase.info/journal/issn1450-569X>

¹² „SCIndeks је српски национални цитатни индекс, развијен да служи као допуна међународним цитатним индексима; реферише домаће часописе категоризоване као периодичне публикације научног карактера. Тренутно садржи 1.009.058 референци из 67.657 чланака од којих 30.924 у виду пуног текста, објављених у 199 различитих домаћих часописа од 2000. у хуманистичким дисциплинама од 1996.а у друштвеним наукама од 1991. Надаље“, преузето 17.03.2016, <http://scindeks.ceon.rs/>

- Google Scholar¹³ (доступни су углавном цитирани радови) – за радове објављене у периоду 2002–2015. године;
- DOAJ¹⁴ – directory of open access journals – за радове објављене у периоду 2002–2015. године.

2.1 Уредници

Деветнаестогодишњи период излагања, од 1997. до 2015. године, обележен је ангажовањем два главна и одговорна уредника. Током 2008. и у првој половини 2009. године, двоброј 17/18 и број 19, имали су гостујуће уреднике. Функција заменика главног и одговорног уредника уведена је 2009. године од броја 20, од истог броја, у ком је дошло и до смене уредника (табела 2).

Променом чланова уредништва, а затим и ступањем на снагу *Акта о уређивању научних часописа* знатно су поопштрени критеријуми уређивања часописа. Од марта 2009. (од бр. 19) сви научни и стручни радови су опремљени научном апаратуром, која укључује апстракт, кључне речи, увођење у проблематику, илустрације и табеле које прате текст, закључне тезе, библиографију цитиране и коришћене литературе. Чланци су редовно категорисани у одговарајуће групе и рецензирани од стране компетентних стручњака са научним звањем. Наведене промене уредништва, поопштрени критеријуми уређивања и стриктна примена Акта о уређењу научних часописа представљају прекретницу у уређивачкој политици што је и разлог зашто је овом анализом обухваћен управо период у распону од 2009. до 2015. године.

¹³ „Google Scholar је, за разлику од WOS-а и Сцопуса, бесплатно доступна база свима који имају приступ Интернету. Ова база обухвата податке о садржају часописа и осталих публикација које издавачи постављају на своје сајтове, као и податке из дигиталних репозиторија, личних веб страница, блогова истакнутих стручњака, препринте итд. Цитати се аутоматски извлаче из текстова који су у отвореном приступу. Највећа предност ове базе је што је бесплатно доступна свима, а највећа мана када су библиометријска истраживања у питању је што је њен опсег непознат. Google не објављује одакле све прикупља податке о цитатима. Пошто се подаци прикупљају аутоматски, дешавају се грешке. Уколико неки рад постоји на мрежи у пуном тексту на више мјеста и у више верзија, он се тако јавља и у овој бази, па није јасно на коју од верзија се односе наведени цитати.” (Raičević, 2013)

¹⁴ „Директоријум часописа отвореног приступа (Directory of Open Access Journals) представља сервис који пружа приступ контролисаним часописима отвореног приступа, односно бесплатним електронским часописима који су задовољили одређене научне и академске критеријуме. Чланци великог броја часописа у директоријуму корисницима су доступни у пуном тексту. Директоријум покрива знатан део научних области, а крајњи циљ је да једног дана све научне дисциплине, на свим значајнијим језицима, буду заступљене“, преузето 21.01.2016. <https://doaj.org/toc/1450-569X>

3. Узорак и методолошки приступ

Уредници	Година	Број часописа
Нада Милашин, главни и одговорни уредник	1997	1, 2
	1998	3, 4
	1999	5
	2000	6
	2001	7
	2002	8
	2003	9
	2004	10, 11
	2005	12
	2006	13/14
	2007	15/16
Нада Милашин, главни и одговорни уредник; Миодраг Вујошевић, Јасна Петрић, гости уредници	2008	17/18
	2009	19
Миодраг Вујошевић, главни и одговорни уредник; Јасна Петрић, заменик главног и одговорног уредника	2009	20, 21
	2010	22, 23
	2011	24, 25, 26
	2012	27, 28
	2013	29, 30
	2014	31, 32
Миодраг Вујошевић, главни и одговорни уредник; Тамара Маричић, заменик главног и одговорног уредника	2015	33, 34

Табела 2. Уредници часописа

Као што је већ наведено, узорком за библиометријску анализу, обухваћени су сви радови објављени од 2009. до 2015. године. У току седмогодишњег периода штампани су бројеви часописа од 19 до 34, у којима је публикувано укупно 145 радова. Подаци неопходни за ово истраживање, прикупљани су индуктивном методом, а библиометријска анализа часописа рађена је *de visu* (са публикацијом у руци), чиме је обезбеђена аутентичност наведених података.

Цитираност часописа је урађена на основу података преузетих из претраживача Google Scholar и цитатних база SCIndeks и Scopus. Број цитата за сваку појединачну свеску преузет је из претраживача Google Scholar и цитатне

базе SCIndeks за период од 2009. до 2015. године, а из цитатне базе Scopus од 2011, када је и почело индексирање *SPATIUM*-а у овој бази, до 2015. године. Како би се лакше уочила разлика у броју цитата у зависности од базе и указло на значај и неопходност консултовања свих доступних извора при евалуацији и рангирању часописа али и вредновању учинка самих аутора, у раду су представљене разлике у анализи цитираности научног часописа *SPATIUM* на примеру рада *Modelling the spatial distribution of Vojvodina's population by using dasymetric method* аутора: Николе Крунића, Вранислава Вајата, Милана Килибарде и Драгутина Тошића објављеног 2011. године у броју 24.

Година	Број часописа	Број објављених радова по свесци	Укупан број радова по годинама
2009	19	7	25
	20	10	
	21	8	
2010	22	8	15
	23	8	
2011	24	10	25
	25	8	
	26	7	
2012	27	7	17
	28	10	
2013	29	12	22
	30	10	
2014	31	13	20
	32	7	
2015	33	13	21
	34	8	
Укупно		145	145

Табела 3. Број радова по свескама

4. Библиометријска анализа

За библиометријску анализу часописа *SPATIUM* коришћени су модели и врсте анализа, као и примери табела из литературе која се бави вредновањем научних часописа и библиометријским истраживањем из различитих научних области. Радом су представљени резултати које се односе на приказан број

радова (Martek and Šute, 2010; Tella and Aisha Olabooye, 2014), категоризација радова (Martek and Šute, 2010), број аутора (Thanuskodi, 2010; Hussain et al., 2011; Jena et al., 2012), афилијација аутора (Thanuskodi, 2010; Hussain et al., 2011; Jokić and Zauder, 2013), заступљеност аутора из различитих држава (Jena et al., 2012), просечан број референци по раду, дужина текстова (Thanuskodi, 2010; Tella and Aisha Olabooye, 2014), укупан број рецензената и њихова међународна заступљеност.

4.1 Број радова по свескама

У наведеном периоду истраживања укупно је објављено 145 радова. Распон броја објављених радова по броју часописа се креће од 7 до 13. Седам радова објављено је у четири наврата у бројевима 19, 23, 26 и 27, а највећи број радова у оквиру једне свеске (13) објављено је у бројевима 31 и 33 (табела 3).

4.2 Категоризација радова

Од 145 радова, колико је објављено у периоду од 2009. до 2015. године, сви радови су рецензирани, а 144 чланка је категоризовано у групу научних односно стручних радова, док један од радова није категоризован. Према *Правилнику* чланци у часописима се разврставају у следеће категорије:

Категорије научних радова у часописима су:

1. оригинални научни рад;
2. прегледни научни рад;
3. кратко или претходно саопштење;
4. научна критика, односно полемика.

Стручни чланци су:

1. стручни рад;
2. информативни прилог;
3. приказ.

Изузетно, у неким областима научни рад у часопису може имати облик монографске студије, као и критичног издања научне грађе.

Од укупног броја чланака, 116 је у категорији научних радова, 27 је стручних радова, 1 рад је категоризован као монографска студија, а 1 чланак није категоризован већ је објављен као технички извештај. У групи научних радова највећи број је прегледних (69), следе оригинални научни радови (30), затим 9 радова у категорији кратко или претходно саопштење и на крају 8 радова у оквиру групе научна критика, односно полемика. У бројевима 26, 27 и 32 објављена су три прилога. Први прилог је информација о међународном

научном пројекту TURaS, други је обавештење о одржавању конференције – RESPAG у организацији Института за архитектуру и урбанизам Србије, а трећи прилог је превод рада *Društveni prostor kao predmet naučno-istraživačkog rada-Spacium*, аутора Милорада Маџуре, објављен 1965. године у часопису *Savremene urbanističke teme*, бр. 2, у издању Института за архитектуру и урбанизам Србије (стр. 43-63)) (табела 4).

SPATIUM	Категоризовани радови						НР*	Прилози**
	ОНР	ПР	КПС	НКП	МС	СР	ТИ	
19 (2009)	4	2			1			
20 (2009)	1	5	1	2		1		
21 (2009)	3	5						
22 (2010)		5	1	1		1		
23 (2010)	1	5	1					
24 (2011)	1	5	2			2		
25 (2011)	1	4		1		2		
26 (2011)	1	3				3		1
27 (2012)		3	1			3		1
28 (2012)	4	6						
29 (2013)	2	4	1	4		1		
30 (2013)	4	3	1			2		
31 (2014)	5	5				2	1	
32 (2014)	1	3				3		1
33 (2015)	2	3	1			7		
34 (2015)		8						
Укупно	30	69	9	8	1	27	1	3
Проценат	20.69	47.59	6.21	5.52	0.68	18.62	0.68	
Укупно	145 радова (100%)							3

Табела 4. Категоризовани и некатегоризовани радови

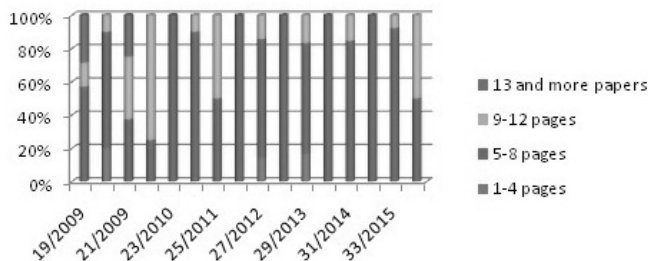
*НР – Некатегоризовани радови;

**Прилози – Прилози (обавештења о пројекту и конференцији, преводи раније објављених радова);

ОНР – Оригиналан научни рад; ПР – Прегледни рад; КПС – Кратко или претходно саопштење; НКП – Научна критика, односно полемика; МС – Монографска студија; СР – Стручни рад; ТИ – Технички извештај

4.3 Дужина радова

На слици 1 приказан је проценат заступљености радова различитих дужина у сваком од анализираних бројева часописа *Spatium*.



Слика 1. Заступљеност радова различитих дужина

Највећи број радова 110 (75,86%) објављен је на 5–8 страна, за четвртину мањи број 26 (17,93%) објављен је на 9–12 страна док је незнатан број текстова штампан на мање од 4 односно више од 13 страна.

4.4 Анализа ауторства

Анализом ауторста за радове објављене у периоду од 2009. до 2014. године може се видети да је једноауторских радова било 48, док је коауторских било укупно 97 (табела 5) од чега је 45 радова са два аутора, 38 радова са три аутора, 9 радова са четири аутора и 8 радова са пет и више аутора (табела 6). Код радова са више аутора преовлађују двоауторски и радови са три аутора, док је број са четири односно пет и више аутора занемарљиво мали. Оно што може да се закључи анализом ауторства је да нема значајнијих осцилација ни у броју једноауторских радова ни у броју коауторских радова од 2010. до 2014. године и њихов однос је приближно исти, али се разлика може уочити код објављених радова у 2009. години и у последња три броја часописа (32, 33 и 34). У бројевима 19, 20 и 21 из 2009. године четири пута је већи број коауторских радова у односу на једноауторске, док је у бројевима 32, 33 и 34, од укупно 28 радова свега три било једноауторских. Број аутора је значајан јер се пун број радова, код вредновања за избор у научна звања, прихвата ако нема више од три коаутора за теоријске радове, односно пет коаутора за нумеричке симулације, односно седам коаутора за експерименталне радове (Ковачевић, 2009).

Ауторство	Година							Укупно	Проценат
	2009	2010	2011	2012	2013	2014	2015		
1 аутор	5	6	12	6	10	8	1	48	33.90
коауторство	20	9	13	11	12	12	20	97	66.90
Укупно	25	15	25	17	22	20	21	145	100.00

Табела 5. Однос једноауторских и коауторских радова

SPATIUM	Број аутора по раду					Укупно
	1 аутор	2 аутора	3 аутора	4 аутора	5 и више аутора	
19 (2009)	3		4			
20 (2009)	1	4	3	2		
21 (2009)	1	2	5			
22 (2010)	3	2	3			
23 (2010)	31	4				
24 (2011)	3	6		1		
25 (2011)	6	1			1	
26 (2011)	3	3	1			
27 (2012)	3	3	1			
28 (2012)	3	4	3			
29 (2013)	6	3	3			
30 (2013)	4	2	2	1	1	
31 (2014)	6	1	4	1	1	
32 (2014)	2	2	2	1		
33 (2015)	1	5	4	1	2	
34 (2015)		3	3	2		
Укупно	48	45	38	9	5	145
Проценат	33.10	31.03	26.21	6.21	3.45	100.00

Табела 6. Број једноауторских и коауторских радова по свесци

4.5 Институцијска припадност аутора

Истраживање везано за афилијацију аутора спроведено је појединачно за три институције: Институт за архитектуру и урбанизам Србије (ИАУС), који је и издавач часописа, Архитектонски факултет (АФ) Универзитета у Београду и Географски факултет (ГФ) Универзитета у Београду на ком се налази Департман за просторно планирање, који су због природе посла самог

Института интересантни за ову студију, док је број аутора осталих институтција из Србије, као и број аутора из иностранства, у табели 7 дат збирно.

SPATIUM	ИАУС	ГФ	АФ	Остале установе у Србији	Установе из иностранства	Појединци (из Србије)	Укупан број аутора из различитих институција
19 (2009)	7	1	4	1	2		15
20 (2009)	1		9	12	4		26
21 (2009)	6			6	8		20
22 (2010)	7		6	1	2		16
23 (2010)	2		2	1	6		11
24 (2011)	4	1	5	3	4	2	19
25 (2011)			5	9	2		16
26 (2011)	5		3	1	2	1	12
27 (2012)			2	4	3	3	12
28 (2012)	9		3	6	2		20
29 (2013)			3	4	14		21
30 (2013)	2		1	12	10		25
31 (2014)	7		4	10	8	1	30
32 (2014)	4		1	10	1		16
33 (2015)	7		1	7	23		38
34 (2015)	5		3	12	3		23
Укупно	66	2	52	99	94	7	320
Проценат	20.63	0.63	16.25	30.94	29.38	2.19	100.00

Табела 7. Институцијска припадност аутора

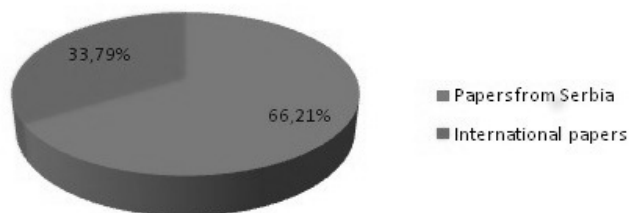
У обрађених 16 свезака радове је објавило укупно 320 аутора од чега је 226 из Србије, а 94 из иностранства. Из самог Института у наведеном периоду је било 66 аутора, са Архитектонског факултета 52 док је са Географског факултета било свега 2 аутора. Укупан број преосталих аутора из Србије, како оних који су запослени у некој од институтција (99) тако и појединаца који нису везани за неку од установа (7), износи 106.

4.6 Број домаћих и иностраних радова

Као што се може видети на слици 2, 96 односно 66,21% је радова домаћих аутора, док је преосталих 49 односно 33,79% радова иностраних аутора¹⁵, што потврђује приличну укљученост међународних радова.

Даље истраживање указује на податак да су радове из иностранства слали аутори из различитих држава света, од којих су најзаступљенији аутори из

¹⁵ Као радови страних аутора узимани су у обзир и коауторски радови домаћих и иностраних аутора



Слика 2. Број домаћих и иностраних радова

Европе али има аутора и из Африке, Азије и Северне и Јужне Америке. Најбројнији су радови аутора из Србије (96), следе радови аутора из Грчке (13), и радови аутора из Словеније, Велике Британије и Немачке (4). Оно што је значајно за сам часопис, његову популарност и углед је чињеница да је број аутора из иностранства у порасту.

4.7 Преглед референци

Попис цитиране и коришћене литературе дат на крају текста један је од основних услова за категоризацију сваког рада. Овакав списак је значајан за квалитет чланака јер пружа увид у грађу коју је аутор користио током свог изучавања, која га је на конкретно истраживање иницирала и подстакла и која је утицала на ток истраживања али је подједнако важан и за оне научне раднике и истраживаче који ће га у будућности користити. Праћењем референци у историјском следу, од коришћене литературе, затим референци у наведеној литератури и тако редом уназад, могуће је сагледати област у којој је научни рад написан и упознати се са конкретним истраживањем (Ђурђевић, 2010).

Највећи утицај на развој конкретне науке имају они научни радови који се највише и цитирају. Највећи број цитата аутори оваквих радова стекли су објављујући значајне научне резултате до којих су током истраживања дошли, а које су њихове колеге препознале као нешто оригинално и подстицајно, нешто што се истиче по свом научном значају и доприносу.

У табели 8 дат је прво број категоризованих радова са референцама по свесци и укупан број категоризованих чланака са референцама за период од 2009. до 2015. године, затим је дат укупан број референци по броју часописа и укупан број референци за све радове објављене током 7 година, а на крају је приказан просечан број реферисане литературе по раду за сваку свеску и на

крају просечан број референци за све радове у часопису *SPATIUM* у назначеном временском раздобљу.

Публикованих 144 научних и стручних радова има укупно 4.158 референци што у просеку износи 28,88 референци по објављеном чланку. Овом анализом није обухваћена просечна старост референци као ни врста реферисане литературе, а није извршено ни искључивање самоцитата аутора што отвара могућности за нека даља истраживања али свакако указује на чињеницу да су аутори у великој мери користили литературу што представља значајан предуслов за озбиљан и свеобухватан истраживачки рад.

	Година																Укупно	
	2009			2010			2011			2012		2013		2014		2015		
	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34		
A*	7	10	8	8	7	10	8	7	7	10	12	10	12	7	13	8	144	
B*	113	163	311	220	142	274	313	191	220	314	408	293	323	197	352	324	4158	
C*	16.14	16.3	38.88	27.5	20.29	27.4	39.13	27.29	31.43	31.4	34	29.3	26.92	28.14	27.08	40.5	28.88	

Табела 8. Преглед референци категоризованих радова

A*Укупан број категоризованих радова са референцама

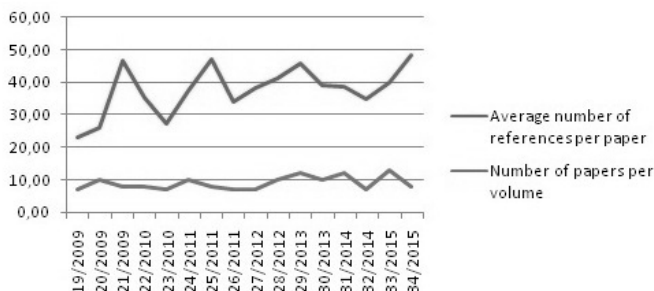
B*Укупан број референци

C*Просечан број референци по раду

Због лакше прегледности на слици 3 је приказан однос броја радова и просечног броја референци за сваку свеску анализираног часописа, у периоду од 2009. до 2015. године. Од броја 21 до броја 33 уочава се готово идентичан образац смањења, односно повећања броја референци у односу на број радова и са незнатним осцилацијама тај број се креће око просечног броја референци израчунатог за цео период (28,88). Значајнија одступања примећују се у бројевима 19 и 20 када је просечан број реферисане литературе био најмањи (16,14 односно 16,3) и у последњој анализираној свесци (34) у којој број реферисане литературе бележи значајан раст, са просечно 40,5 референци по раду.

4.8 Анализа кључних речи

За сажето представљање садржаја научних радова од значаја је правилно дефинисање кључних речи (Јевремов, 2009). Оне представљају основ за класификацију докумената и њихово индексирање у различитим базама података.



Слика 3. Однос броја радова и просечног броја референци по свесци

Узорак за анализу кључних речи су, и у овом случају, били радови објављени у периоду од 2009. до 2015. године. Од укупно 145 публикованих радова у 3 рада нису изведене кључне речи (један раду у броју 20 из 2009. и по један рад у бројевима 22 и 23 из 2010. године), што представља висок степен заступљености кључних речи у наведеном корпусу (Табела 9). У преостала 142 рада укупно је пронађено 673 кључне речи. Од овог броја 627 је само једном заступљено док преосталих 46 речи има различит број понављања. 28 речи понавља се два пут, 9 речи три пута, 3 речи четири пута, код 4 речи било је пет понављања, док се код две речи са највише понављања 1 јавља осам пута, а 1 кључна реч десет пута. Две најчешће понављане кључне речи су: *Србија* (10) и *одрживи развој* (8). *Животна средина*, *идентитет*, *просторно планирање* и *одрживост* су четири израза која су по пет пута употребљена као кључне речи. А са три понављања следе *Београд*, *Климатске промене* и *планирање* (Слика 4).

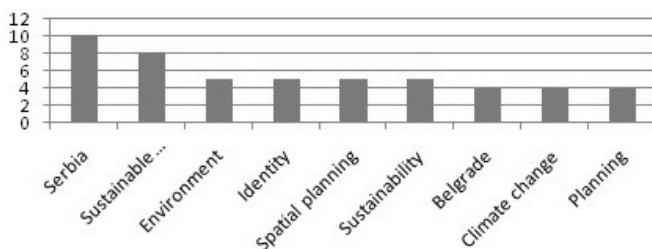
У свим радовима кључне речи су одређивали сами аутори и с обзиром на то да не постоји контролисани речник из области просторног планирања уочава се нестандаризованост коришћења термина.

4.9 Рецензенти

Квалитет рада и потврду презентованог научног резултата мора, пре објављивања, да потврди неко ко је компетентан за ту научну област. Уредништво *SPATIUM-a* води рачуна да израду рецензија повери најеминентнијим стручњацима из земље и иностранства, поштујући, при том, узајамну анонимност и аутора и рацензената. Сталним залагањем главног и одговорног уредника, као и заменика главног уредника, да привуку што већи број страних рецензената доприносе вишем статусу часописа. У табели 10 приказан је број рецензената из Србије и иностранства за сваки од педесет

Година	Број кључних речи	Број радова са кључним речима	Радови без кључних речи
2009	125	24	1 (20)
2010	66	13	1 (22)
2011	117	25	1 (23)
2012	73	17	
2013	104	22	
2014	90	20	
2015	98	21	
Укупно	673	142	3

Табела 9. Кључне речи



Слика 4. Најчешће понављане кључне речи

бројева обухваћених овим истраживањем, као и укупан број домаћих и страних рецензената у периоду од 2009. до 2015. године.

5. Цитатна анализа

Један од основних критеријума вредновања научног рада је цитираност објављених радова у којима је представљен крајњи резултат одређене фазе научних истраживања. Вредновање на основу цитираности може се применити на појединачне научнике, на универзитете и научне институције у целини и на часописе у којима се научни радови објављују (Raičević, 2013).

Ова врста вредновања може се применити и на сваку појединачну научну област. Управо се стварање цитатног индекса јавља као последица жеље и

SPATIUM	Година																Укупно	Проценат
	2009				2010				2011				2012					
	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34		
Рецензенти из Србије	7	11	11	11	8	17	9	11	8	15	17	15	23	14	18	7	202	71,13
Рецензенти из иностранства	0	6	1	4	5	2	6	1	5	6	3	4	6	4	14	15	82	28,87
Укупно	7	17	12	15	13	19	15	12	13	21	20	19	29	18	32	22	284	100

Табела 10. Рецензенти

потребе да се формира база података која ће омогућити праћење научног истраживања, праћење научне идеје као и напредовања и развоја конкретне науке. Сваки научни радник током свог истраживања користи одређену литературу. Навођењем тих радова везује своје резултате истраживања са резултатима који су раније објављени (Filipi Matutinović, 2013).

Напредовање истраживача и њихов избор у виша научна звања умногоме зависе управо од броја цитата њихових радова. И поред тога што постоје недостаци који се јављају приликом евалуације научног рада на основу података о цитираности, као што су велики број самоцитата¹⁶ и коцитата, манипулација цитатима у оквиру научних кругова, непостојање адекватне везе између посматраног и цитираног рада, мотиви цитирања или нецитирања који не морају по правилу бити научни и др, постоје и позитивне стране јер велика цитираност одређеног аутора свакако представља признање да је његов рад имао велики утицај на потоња истраживања, затим цитати су јавно доступни и проверљиви, њихово прикупљање и анализа су олакшани цитатним базама и претраживачима.

¹⁶ Код бројања цитата важан је однос између цитата и самоцитата. Податак да је број самоцитата већи од уобичајених 20% у односу на укупан број цитата може да укаже на покушај манипулације. „Висока стопа самоцитираности карактеристична је за водеће научне часописе како због сталног високог квалитета радова објављених у њима, тако и због јединствене или оригиналне тематске профилације... Међутим, има часописа који имају тако високу стопу самоцитираности да оне деформишу укупну представу о цитираности таквих часописа...у случају знатнијег одступања...ISI проверара да ли је реч о самоцитирању које се врши да би се вештачки подигао фактор утицајности часописа.” (Vukasović, 2009)

5.1 Извори података за цитираност радова у часопису *SPATIUM*

Подаци о цитираности радова објављених у часопису *SPATIUM* преузети су из следећих цитатних база и претраживача:

- SCIndeks (Српски цитатни индекс), цитатна база и претраживач научних часописа објављених у Србији;
- Google Scholar, цитатни претраживач компаније Google;
- Scopus, цитатна база међународне издавачке компаније Elsevier.

Временски распон у оквиру ког су доступни подаци о цитираности за радове у часопису *SPATIUM*:

- Српског цитатног индекса – за радове објављене у периоду од 2000. до 2013. године;
- Google Scholar – за радове објављене у периоду од 2000. до 2015. године;
- Scopus¹⁷ – за радове објављене у периоду од 2011. до 2015. године.

У табели 11 дати су подаци о цитираности часописа *SPATIUM* од 2009. до 2015. године, од броја 19 до броја 34. Укупан број цитата за SCIndeks и Google scholar, прво је дат за период од 2009. до 2015. године, а у загради и за период од 2011. до 2015. који се поклапа са укупном цитираношћу у цитатној бази Scopus како би могло да се направи поређење сва три извора цитираности. Може се уочити да сваки од наведених извора цитираности даје различите податке о броју цитата за конкретан број часописа у оквиру одређене године, а самим тим се разликује и број цитата појединачног рада у оквиру сваког броја. Зато је важно да се за научника из Србије цитатна анализа врши на основу свих доступних извора и да се никада не врши поређење добијених резултата из једне базе за неког научника са добијеним резултатима из друге базе за неког другог научника (Filipi Matutinović, 2013).

5.2 Разлике у приказу цитираности на примеру једног рада

На примеру рада *Modelling the spatial distribution of Vojvodina's population by using dasymetric method* аутора: Николе Крунића, Вранислава Вајата, Милана Килибарде и Драгутина Тошића биће приказана разлика у томе колико је пута рад цитиран, где је цитиран и када према подацима преузетим из три различита извора. Рад је објављен 2011. године у броју 24. Подаци су из све три базе: SCIndeks, Google Scholar и Scopus преузети истог дана, 19. јануара. 2016. године.

¹⁷ Ова база података даје могућност ауторима да виде цитираност својих радова, помоћу опције Citation tracker. Веома је битно да постоји могућност искључивања самоцитата помоћу опције exclude autocitations. Подаци о цитирању одређеног чланка дати су и у EXCEL табели, која показује број цитата за дати чланак за сваку годину посебно, почевши од 1996, као и укупан број цитата.

SPATIUM	Година																Укупно
	2009		2010		2011		2012		2013		2014		2015				
	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	
Укупан број радова	7	10	8	8	7	10	8	7	7	10	12	10	12+1	7	13	8	145
SCIndex	12	8	23	10	14	19	8	8	3	1	1	0	-	-	-	-	107 (40)
Scopus	-	-	-	-	-	13	13	9	8	8	4	9	4	5	0	0	73 (73)
Google Scholar	17	31	62	34	36	32	18	15	16	12	10	9	4	2	1	-	299 (119)

Табела 11. Преглед цитираности

Приказ цитираности у цитатној бази SCIndeks Како се може видети рад је у цитатној бази SCIndeks цитиран три пута. Године 2011. цитиран је у *Zborniku radova Geografskog fakulteta Univerziteta u Beogradu*, 2012. у часопису *SPATIUM* и 2014. године у часопису *Geonauka*. Сва три рада приказана су и као резултати у претраживачу Google Scholar, а у цитатној бази Scopus као резултат, од горе наведених, приказан је само рад *Spatial-functional organization of settlements in Vojvodina*, из 2012. године. У сва три случаја рад је цитиран од стране домаћих аутора, од којих је један цитат, а два су самоцитати (слика 5)¹⁸.

Приказ цитираности у претраживачу Google Scholar У претраживачу Google Scholar стоји да је исти рад цитиран 9 пута. Радови нису дати у историјском следу нити су поређани алфабетски, већ, вероватно, оним редоследм којим је цитатни претраживач наилазио на податке на светској мрежи. Радови су објављени у периоду од 2011. до 2015. године. Према Google Scholar-у, рад *Modelling the spatial distribution of Vojvodina's population by using dasymetric method* цитиран је 2011. године два пута, једном у зборнику радова, а једном у часопису, 2012. два пута у часопису и једном у зборнику радова, 2014. године два пута у радовима објављеним у часописима, а 2015. године по јаданпут у докторској дисертацији и поглављу у монографској публикацији. Међу ових девет радова налазе се и сва три рада приказана као резултати у цитатној бази SCIndeks. Од укупног броја цитата, рад *Modelling the spatial distribution of Vojvodina's population by using dasymetric method* цитиран је, у претраживачу Google Scholar, од стране домаћих аутора 7 пута, а два пута су га цитирали аутори из иностранства, од тога је 6 самоцитата и 3 цитата (слика 6)¹⁹.

¹⁸ <http://scindeks.ceon.rs/Related.aspx?artcit=1450-569X1124045K>
¹⁹ http://scholar.google.com/scholar?cites=1285848406558262205&as_sdt=2005&sciodt=0,5&hl=sr



Слика 5. Приказ цитираности у цитатној бази SCIndexs

Приказ цитираности у цитатној бази Scopus Увидом у резултате добијене из цитатне базе Scopus можемо видети да је рад *Modelling the spatial distribution of Vojvodina's population by using dasymetric method*, у овој бази цитиран четири пута, по једном у радовима објављеним 2012, 2013, 2014. и 2015. године. Од тог броја, у 3 случаја ради се о радовима домаћих аутора и сва три су самоцитати, а четврти је рад аутора из иностранства и у питању је hetero citat (слика 7)²⁰.

Даљом анализом видимо да је рад *Spatial-functional organization of settlements in Vojvodina*, из 2012. године, као резултат претраге, наведен у све три цитатне базе. Рад под називом *Dasymetric modelling of population dynamic in urban areas*, из 2013. године, је као резултат добијен само у цитатној бази Scopus. А радови *A fine-scale spatial population distribution on the High-resolution Gridded Population Surface and application in Alachua County, Florida* i *Dasymetric Mapping of Population Distribution in Serbia Based on Soil Sealing Degrees Layer* као резултат су добијени у цитатној бази Scopus и претраживачу Google Scholar.

Посматрајући ове примере може се закључити да би за евалуацију часописа и његово рангирање у одговарајућу категорију, било неопходно консултовати добијене резултате из свих доступних извора, како ни сам часопис, а ни аутори који у њему објављују не би били оштећени неадекватном категоризацијом.

²⁰ <http://www.scopus.com/results>



Modelling the spatial distribution of Vojvodina's population by using dasymetric

Q

Академик

Modelling the spatial distribution of Vojvodina's population by using dasymetric method
N Krunić, B Bajat, M Kilibarda, D Tošić - Spatium, 2011 - doi:serbia.nb.rs
Cartographic presentation of heterogeneity/homogeneity in the spatial distribution of population is still a major problem in modern geography, and other geo-sciences as well. The traditional method of thematic or choropleth mapping rarely gives satisfactory results. ...
9 пута наведен Сродни чланци Све верзије (6) Цитирај Сачувај Више

plotgooglemaps: The r-based web-mapping tool for thematic spatial data
M Kilibarda, B Bajat - Geomatica, 2012 - pubs.cig-acsg.ca
Google Maps are increasingly used for communication throughout many map-based services and maps, embedded on third-party websites via the Google Maps API. The main objective of this paper is to present a solution for an easy creation of an interactive web ...
13 пута наведен Сродни чланци Све верзије (5) Цитирај Сачувај Више

[PDF]

Dasymetric mapping of spatial distribution of population in Timok Region

B Bajat, N Krunić, M Kilibarda - Proceedings of international ..., 2011 - e-science.amres.ac.rs

Summary: Dasymetric mapping of population distribution represents very functional visualization method used in spatial demographic analysis. The main advantage of dasymetric mapping over standardized cartographic methods (choropleth maps) used for ...

5 пута наведен Сродни чланци Све верзије (2) Цитирај Сачувај Више

[PDF]

Spatial modelling of population concentration using geographically weighted regression method

B Bajat, N Krunić, M Kilibarda - Journal of the ..., 2011 - doi:serbia.nb.rs

This paper presents possibilities of applying the geographically weighted regression method in mapping population change index. During the last decade, this contemporary spatial modeling method has been increasingly used in geographical analyses. On the example ...

2 пута наведен Сродни чланци Све верзије (10) Цитирај Сачувај Више

[PDF]

Laboratory for development of open source geospatial technologies—role in education and research

M Kilibarda, B Bajat, N Branislavić - Geonauka, 2014 - researchgate.net

Abstract. International Cartographic CBOs (International Cartographic Association-ICA) in partnership with the Open Source Geospatial Foundation-OSGeo has started the initiative ICA-OSGeo Labs to promote and use open source technologies in education and ...

Сродни чланци Све верзије (5) Цитирај Сачувај Више

[PDF]

MODELAGEM ESPACIAL DINÂMICA DOS DETERMINANTES SOCIAIS E AMBIENTAIS DA MALÁRIA E SIMULAÇÃO DE CENÁRIOS 2020 PARA MUNICÍPIO ...

TD do Curso - mtc-m21b.sid.inpe.br

RESUMO Em pleno século XXI a malária continua sendo uma das endemias de maior magnitude no mundo. Segundo a Organização Mundial da Saúde, no ano de 2013 ocorreram 132 milhões de novos casos, concentrados em países periféricos, sobretudo, ...

Цитирај Сачувај Више

Spatial-functional organization of settlements in Vojvodina

N Krunić - Spatium, 2012 - doi:serbia.nb.rs

This paper summarizes the results of recent exploration of spatial and functional organization of Autonomous Province of Vojvodina in the Republic of Serbia (hereinafter referred to as "Vojvodina") based on identification of the level of development of spatial ...

1 пута наведен Сродни чланци Све верзије (5) Цитирај Сачувај Више

Dasymetric Mapping of Population Distribution in Serbia Based on Soil Sealing Degrees Layer

N Krunić, B Bajat, M Kilibarda - Surface Models for Geosciences, 2015 - Springer

Abstract This paper outlines a methodology used to disaggregate a census population in order to more accurately determine the population distribution over a regional area or a state scale. Data regarding population distributions are usually accessible at the level of ...

1 пута наведен Сродни чланци Све верзије (3) Цитирај Сачувај Више

[HTML]

A fine-scale spatial population distribution on the High-resolution Gridded Population Surface and application in Alachua County, Florida

P Jia, Y Qiu, A E Gaughan - Applied Geography, 2014 - Elsevier

Abstract Geospatial techniques, using Geographic Information Systems and remote sensing data, have become more commonly used with dasymetric modeling of fine-scale demographic data. In this study, we apply a dasymetric approach using the Heuristic ...

4 пута наведен Сродни чланци Све верзије (2) Web of Science: 2 Цитирај Сачувај Више

[PDF]

Промене у дистрибуцији руралног становништва Србије

В Глигоријевић, М Девеџић, V Gligorjević - zbornik.geog.bg.ac.rs

Вера Глигоријевић¹, Мирјана Девеџић² Универзитет у Београду–Географски факултет Новод: Циљ овог рада је допринес дискусии о урбано-руралној дистрибуцији становништва у Србији. У контексту европске регионалне политике, рад се посебно ...

Сродни чланци Цитирај Сачувај Више

Слика 6. Приказ цитираности у претраживачу Google Scholar

94

Инфотека, год. 16, бр. 1–2, август 2016.

Scopus			
☐ 1	Modelling the spatial distribution of Vojvodina's population by using dasymetric method	Krunić, N., Bajat, B., Kilibarda, M., Tošić, D.	2011 Spatium (24), pp. 45-50 4 Cited by
 View at Publisher Show abstract Related documents			
Export Download View citation overview View Cited by More...			
☐ 1	Dasymetric mapping of population distribution in serbia based on soil sealing degrees layer	Krunić, N., Bajat, B., Kilibarda, M.	2015 Lecture Notes in Geoinformation and Cartography 211, pp. 137-149
 View at Publisher Show abstract Related documents			
☐ 2	A fine-scale spatial population distribution on the High-resolution Gridded Population Surface and application in Alachua County, Florida	Jia, P., Qiu, Y., Gaughan, A.E.	2014 Applied Geography
 View at Publisher			
☐ 3	Dasimetrično modeliranje dinamike prebivalstva na urbanih območjih [Dasymetric modelling of population dynamics in urban areas]	Bajat, B., Krunić, N., Samardžić Petrović, M., Kilibarda, M.	2013 Geodetski Vestnik
			
☐ 4	Spatial-functional organization of settlements in Vojvodina	Krunić, N.	2012 Spatium
 View at Publisher			

Слика 7. Приказ цитираности у цитатној бази Scopus

6. Категоризација домаћих научних часописа за област Саобраћај, урбанизам и грађевинарство

Сагласно члану 27. став 1. тачка 4) и члану 25. став 1. тачка 5) Закона о научноистраживачкој делатности,²¹ Министарство просвете, науке и технолошког развоја, након размотрених приговора на категоризацију домаћих научних часописа за 2013. годину, усвојило је коначну листу категоризованих научних часописа за 2013. годину. На листи часописа за област Саобраћај, урбанизам и грађевинарство за 2013. годину часопис SPATIUM се налази на првом месту у категорији M24. Листа из 2013. године је последња званична листа категоризованих домаћих научних часописа, које је Министарство усвојило. Најновији предлог категоризације за 2014. годину постављен је на сајту Министарства просвете, науке и технолошког развоја. И на последњој, прелиминарној листи часопис SPATIUM налази се у истој категорији M24.

²¹ Закон о научноистраживачкој делатности („Службени гласник РС”, бр. 110/05, 50/06-испр. и 18/10)

7. Закључак

Часопис *SPATIUM* излази већ 19 година. Овом анализом је обухваћен само кратак период од 2009. до 2015. године који се отприлике поклапа са променом уредништва и доношењем *Акта о уређивању научних часописа* 2009. године од стране Министарства за науку и технолошки развој Републике Србије. Циљ истраживања је био да се прикаже улога и значај часописа *SPATIUM*, његов мултидисциплинарни приступ истраживањима и објављивању радова из различитих научних области.

На основу анализе објављених текстова види се да је број научних радова четири пута већи од броја стручних чланака. Од укупног броја радова проценат објављених научних радова је 80% док је стручних 18,62 из чега проистиче да је *SPATIUM* пре свега усмерен ка научној комуникацији, мада никако не треба занемарити ни његову улогу у размени стручних искустава.

Сталном борбом за квалитет и утицајност, уредништво *SPATIUM-a* доприноси вишем статусу часописа. Циљ уредништва је да се часопис у наредном периоду пријави за поступак евалуације ради његовог уврштавања на Thomson Reuters листу. Евентуалним уласком на ову листу, часопис би остварио још већу афирмацију у међународним научним и стручним круговима за област просторног и урбанистичког планирања, архитектонског пројектовања и других сродних тема чиме би се повећао и фактор утицајности *SPATIUM-a* као квантитативни исказ успешности часописа.

Литература

- Антонић, Сања, Кринуловић, Оја и Каваја Станишић, Дејана. “Анализа цитираности: Тренутно стање и перспективе”. *Панчевачко читалиште*, год. VIII, бр. 15 (2009): 9–11, приступљено 19. I 2016, http://citaliste.rs/casopis/br15/antonic_sanja_1.htm.
- Ђурђевић, Душан Ж. “Citation analysis of the University of Priština located in Kosovska Mitrovica”. *Sociologija*, vol. 52, number 4 (2010): 421–446, приступљено 19. I 2016, <http://www.doiserbia.nb.rs/img/doi/0038-0318/2010/0038-03181004421D.pdf>, DOI10.2298/S0C1004421D, accessed 2016-01-19.
- Филипи Матутиновић, Стела. *Научне информације у Србији : протокол, доступност, вредновање*. Приступљено 19. I 2016, <http://kobson.nb.rs/upload/documents/oNamaPredavanja/PR2013TekstZaDoktorante.pdf>, Belgrade, 2013.
- Hussain, Akhtar, Fatima, Nishat and Kumar, Devendra. “Bibliometric analysis of the ‘electronic library’ journal (2000-2010)”. *Webology*, vol. 8, number 1 (2011): 87, приступљено 19. I 2016, <http://www.webology.org/2011/v8n1/a87.html>.

- Jena, Kamal Lochan, Swain, Dillip K. and Sahu, Sada Bihari. "Scholarly communication of the electronic library from 2003-2009: a bibliometric study". *The Electronic Library*, vol. 30, number 1 (2012): 103–119, приступљено 17. II 2016, <http://www.emeraldinsight.com/doi/pdfplus/10.1108/02640471211204097>.
- Jevremov, Tanja. "Informatička vrednost autorskih ključnih reči u člancima predstavljanim u Srpskom citatnom indeksu". In *Scientific and Professional Conference SNTPI '09*, pages 91–94, 2009. <http://ceon.rs/pdf/jevremov.pdf>.
- Jokić, Maja and Zauder, Krešimir. "Bibliometric analysis of rural sociology/sociology and space from 1963-2012". *Sociologija i prostor*, vol. 196, number 2 (2013): 331–349, приступљено 17. I 2016, http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=161510,DOI:10.5673/sip.51.2.11.
- Ковачевић, Милош. "Категоризација научних часописа: Активност Матичног научног одбора за језик и књижевност (за хуманистичке науке)". *Панчевачко читалиште*, год. VIII, бр. 15 (2009): 4–8, приступљено 19. I 2016, http://citaliste.rs/casopis/br15/kovacevic_milos_1.htm.
- Martek, Alisa and Šute, Snježana. "Bibliometric analysis of the journal arhivski vjesnik in the period 2001-2009". *Arhivski vjesnik*, number 53 (2010): 165–176, приступљено 19. I 2016, <http://hrcak.srce.hr/62500>.
- Raičević, Saša. "Database as a source of information on the evaluation of scientific papers in medicine". *Medical Journal of Montenegro*, vol. 1, number 15 (2013): 9–11, 2013. приступљено 21. I 2016, <http://mjmn.me/11/DatabaseAsASource/DatabaseAsASource.html,doi:10.5937/cma1.2822>.
- Tella, Adeyinka and Aisha Olabooye, Ayotola. "Bibliometric analysis of african journal of library, archives and information science from 2000-2012". *Library Review*, vol. 63, number 4/5 (2014): 305–323, приступљено 19. I 2016, <http://www.emeraldinsight.com/doi/pdfplus/10.1108/LR-07-2013-0094,DOIhttp://dx.doi.org/10.1108/LR-07-2013-0094>.
- Thanuskodi, S. "Bibliometric analysis of the journal library philosophy and practice from 2005-2009". *Library Philosophy and Practice*, pages 1–6, 2010. Приступљено 19. I 2016, <http://unllib.unl.edu/LPP/thanuskodi-lpp.htm>.
- Tóth, Tibor. "Sastanak s Jacekom Ciesielskim, direktorom izdavačke kuće Versita iz Varšave". *HID-DROBILICA*, vol. IV, number 2 (2007): 2–3, приступљено 15. II 2016, http://hidd.hr/wp-content/uploads/sites/178/2013/10/Drobilica_br2_2007.pdf.
- Вучковић, Жељко. "Функције часописа у развоју научне комуникације". *Панчевачко читалиште*, год. VIII, бр. 15 (2009): 2–3, приступљено 14. II 2016, http://www.citaliste.com/casopis/br15/Citaliste_br15.pdf.
- Vukasović, Predrag. "Analysis of citations, origins, purposes and basic principles". *Strani pravni život*, number 1 (2009): 211–230.

Web странице

“DOI Srbija - Naučni časopisi objavljeni u Srbiji u punom tekstu” The European Library, преузето 15.02.2016, <http://www.theeuropeanlibrary.org/tel4/collection/a0155?id=a0155&locale=sr>

“Kategorizacija časopisa” Ministarstvo prosvete, nauke i tehnološkog razvoja, преузето 14.02.2016, <http://www.mpn.gov.rs/casopisi/>

SCIndeks - Srpski citatni indeks, преузето 15.02.2016, <http://scindeks.ceon.rs>

“Spatium” doiSerbia, преузето 15.02.2016, <http://www.doiserbia.nb.rs/journal.aspx?issn=1450-569X>

“Spatium” SCIndeks - Srpski citatni indeks, преузето 15.02.2016, <http://scindeks.ceon.rs/journaldetails.aspx?issn=1450-569X>

“Spatium, 1450-569X” Academic Journals Database - Disseminating quality controlled scientific knowledge, преузето 15.02.2016, <http://journaldatabase.info/journal/issn1450-569X>

“Spatium, 1450-569X (Print)” Directory of open access journals, преузето 15.02.2016, <http://doaj.org/toc/d3f460f04f2f4f849aa9feef69cec258>

“SPATIUM-The Journal of Institute of Architecture and Urban & Spatial Planning of Serbia” de Gruyter Open, преузето 15.02.2016, <http://www.degruyter.com/view/j/spat>

<http://scindeks.ceon.rs/Related.aspx?artcit=1450-569X1124045K>

http://scholar.google.com/scholar?cites=1285848406558262205&as_sdt=2005&scioldt=0,5&hl=sr

Електронске књиге и нова димензија читања

УДК 004.738.5:655.41

САЖЕТАК: Развој технологије допринео је томе да се број књига у електронском формату свакодневно увећава. Приступ електронским књигама омогућавају, бесплатно или уз новчану надокнаду, издавачи и дистрибутери електронских издања. У Србији постоји неколико издавача који се баве издавањем књига у електронском формату, а приметан је и пораст дигиталних библиотека које омогућавају приступ електронским књигама. Електронске књиге могу се читати на рачунарима, лаптоповима, таблетима и паметним телефонима, али су за дуже читање најпогоднији читачи електронских књига. Једноставан приступ, двадесетчетворочасовна доступност, економичност у смислу цене и простора и друге предности електронских књига доприносе повећању читалачке публике.

КЉУЧНЕ РЕЧИ: дигитализација, електронска књига, електронско издаваштво, уређаји за читање електронских књига, читање.

РАД ПРИМЉЕН: 25. март 2016.

РАД ПРИХВАЋЕН: 07. мај 2016.

Гордана Недељков

gordana.nedeljkov06@gmail.com

Андрићев институт, Вишеград

1. Увод

Књига је током свог вишевековног трајања мењала свој физички облик, али не и сврху. Одувек је служила као инструмент за бележење, чување и преношење знања и информација. Према дефиницији Унеска књига је „непериодична штампана публикација од најмање 49 страна које не подразумевају корице“.¹ У Енциклопедији Британика наведено је да књига представља инструмент комуникације, при чему се књига посматра као писана или штампана порука одређене дужине, намењена јавном коришћењу и забележена на материјалу који је довољно издржљив да омогући лако преношење.²

¹ Business dictionary, <http://www.businessdictionary.com/unhbox/voidb@x\hbox{definition}/book.html>, приступљено 14. IV 2016.

² Encyclopedia Britannica, <http://www.britannica.com/EBchecked/topic/73295/book>, приступљено 21. III 2016.

Књиге су данас најчешће заступљене у штампаном облику, али се све чешће појављују и у дигиталном формату. Електронске књиге најрадије прихватају нове генерације, али и старији корисници, нарочито они којима коришћење рачунара и нових информационих технологија није страно, све чешће показују интересовање за читањем књига у електронском облику.

Како се може приступити електронским књигама, каква је разлика у читању електронских и штампаних формата књиге и које су предности и мане електронских у односу на штампана издања, нека су од питања којима ћемо се бавити у овом раду. Осим тога, у раду ће бити речи о електронском издаваштву и дигиталним библиотекама у Србији и у свету, о уређајима за читање електронских књига и променама које су настале у самом процесу читања.

2. Појам и настанак електронских књига

Под електронским књигама подразумевају се монографске публикације које постоје истовремено на папиру и у дигиталној форми или само у дигиталној форми. Развој технологије допринео је томе да се број електронских књига свакодневно увећава. Електронске књиге се најчешће дистрибуирају на компакт дисковима и преко мреже, а читају се на рачунару, мобилним уређајима или тзв. читачима електронских књига. Првобитна замисао постојања књига у електронском облику била је да се помоћу њих омогући остављање коментара на прочитани текст и комуникација са аутором (Васиљевић, 2006). Ипак, како наводи ауторка горепоменутог рада, ова идеја није заживела, а електронске књиге наставиле су да постоје имајући у суштини исту функцију као и књиге у штампаној форми – као средство за преношење информација и знања, али сада на бржи и једноставнији начин. Са развојем електронских књига дошло је до промене у самом процесу читања, а позитивни аспекти тих промена препознати су у литератури и потврђени у пракси што је допринело повећању популарности књига у електронском облику.

У терминологији енглеског језика постоје два појма која се односе на дигиталне књиге и који објашњавају како су настале: „born-digital“ и „made-digital“.³ Термин „born-digital book“ односи се на књиге које су оригинално настале у дигиталном облику. Прва књига која је објављена само у електронском облику била је роман „Riding the Bullet“ америчког аутора Стивена Кинга, објављен у марту 2000. године (Трифунковић, 2005). Други термин, „made-digital“, подразумева књиге које су из штампане верзије претворене у

³ У преводу на српски IFLA/UNESCO Манифеста за дигиталне библиотеке користе се термини „изворно дигитално рођени“ и „дигитализовани“ садржај.

електронску. Процес дигитализације, односно „преношење књиге из штампаног или рукописног у машински читљив облик уз употребу нарочитих хардверских и софтверских решења“ (Вранеш и Марковић, 2008), подразумева сложени скуп послова који обухватају одабир грађе, саму дигитализацију штампаног издања, обраду и обележавање текста да би се учинио претраживим и чување књиге на неком медијуму или на серверу у одређеном формату. Поред општих формата за документа у електронском облику, као што су ASCII, HTML и PDF, који могу да се користе и за електронске књиге, специјализовани формати књига у електронском облику су ePub, Mobi, Nook и други.

3. Како до електронских књига?

3.1 Дигиталне библиотеке и издавање електронских књига

Књиге у дигиталном облику могу се наћи као део пројеката дигитализације, у дигиталним библиотекама широм света. Дигитална библиотека је збирка дигиталних објеката потврђеног квалитета, који су, уз употребу нових технологија, створени или прикупљени и организовани у складу са међународним принципима и стављени на располагање корисницима (IFLA/UNESCO, 2010). Већи део публикација у дигиталним библиотекама налази се у отвореном приступу, који подразумева да „сваки корисник који има право приступа интернету има право да чита, преузима, чува, штампа и користи дигитални садржај радова објављених у отвореном приступу, са једином обавезом да их коректно цитира“ (Филипи-Матутиновић, 2013). У свету је пракса да аутор потписивањем лиценце даје пристанак на отворени приступ своме делу. Слободан приступ електронским књигама у пуном тексту омогућавају OA (Open Access) базе података, а неке од њих су: Project Gutenberg, Google Books, Europeana, Open Library, Internet archive, Digital Public Library of America, Daily Lit, Classic Reader и National Academic Press.

Приступ електронским књигама омогућавају, такође, издавачи и дистрибутери електронских издања, бесплатно или уз новчану надокнаду. Електронске књиге се набављају путем претплата или путем лиценци било да се ради о једном наслову или о пакету. Од комерцијалних добављача електронских књига најпознатији су: Questia, netLibrary, ebrary, Oxford Reference Online, Springer, Elsevier, ScienceDirect, EBSCO и Emerald.

Ниво доступности и услови приступа електронским књигама зависе од добављача и од ауторских права. Приликом издавања електронских књига мора се водити рачуна о поштовању ауторских права, као и када се ради о класичном штампаном издању. Издавачке куће специјализоване за издавање електронских

књига често објављују управо она издања којима су истекла ауторска права. Уколико то није случај, издавање електронских књига подложно је закону о ауторском праву земље на чијој територији се публикација издаје. Већина земаља пружа заштиту и страним ауторима, односно носиоцима ауторских права, уколико они долазе из земље која је потписница неког од међународних уговора у области ауторских права. Један од таквих уговора је Бернска конвенција, глобални споразум о заштити ауторских права на интернету, према којој национални закони сваке земље подразумевају аутоматску заштиту за ауторска дела, у складу са законом земље потписнице. Србија је једна од земаља потписница Бернске конвенције, а о ауторским правима у Србији се додатно може информисати у Закону о ауторском и сродним правима.

3.2 Електронске књиге у свету

Први пројекат који је омогућио приступ књигама у електронском формату представља Пројекат Гутенберг,⁴ са мисијом да подстиче стварање и дистрибуцију електронских књига. Покретач овог пројекта био је Мајкл Харт који је још 1971. године почео да дигитализује текстове који не подлежу ауторским правима и да их прво поставља на сервере, а касније и на интернет. Временом се број дигитализованих књига повећавао, да би данас на сајту пројекта било постављено преко 50.000 електронских књига. Све књиге које су укључене у Пројекат Гутенберг су у пуном тексту и у потпуности претраживе. На почетној страници Пројекта Гутенберг наводи се да је коришћење дигитализованих књига у складу са Законом о ауторским правима у Сједињеним Америчким Државама, а да су корисници из других земаља у обавези да провере закон о ауторским правима у својој земљи пре преузимања и дистрибуције књига.

Пројекат дигитализације европске културне баштине под називом Еуропеана⁵ покренут је са идејом да се омогући слободни приступ дигитализованом садржају, заштита културног блага и размена идеја и информација. Преко портала Еуропеане могућ је приступ преко 24 милиона докумената у дигиталном облику, међу њима и електронским књигама. Број партнера и количина доступног садржаја на порталу Еуропеана свакодневно расте.

Google books⁶ омогућава претраживање и приступ књигама у електронском облику. Видљивост добијених резултата претраживања и доступност зависе од ауторских права. Уколико су ауторска права истекла или су дата компанији

⁴ Пројекат Гутенберг, <http://www.gutenberg.org/>, приступљено 23. III 2016.

⁵ Еуропеана, <http://www.europeana.eu/portal/>, приступљено 14. IV 2016.

⁶ Google books, <https://books.google.com/>, приступљено 11. IV 2016.

Google books и књига се налази на интернету као јавно добро, могуће је преузимање књиге у целости. У другим случајевима, аутор или други носилац ауторских права одређује ниво доступности публикације. У оквиру овог пројекта постоји сервис Google Books Library Project који као резултат претраге даје библиографске податке о књизи и, уколико је књига заштићена ауторским правима и недоступна за преузимање, везе ка интернет продавницама где се књига може купити и библиотекарима у којима се књига може позајмити.

Open Library⁷ представља базу података у којој се налазе везе које воде до више од милион електронских књига. Одређени број књига је у слободном приступу, а постоји и опција позајмљивања књига из онлајн библиотеке.

Поменуте базе података садрже тематски разноврсне електронске књиге, те се поред дела класичне и савремене књижевности у њима могу наћи и књиге из различитих научних области. На тај начин, постављањем дигиталног садржаја на мрежу, корисницима широм света је омогућено да брзо и једноставно дођу до потребне литературе коју ће користити за учење, истраживање или читање из забаве.

3.3 Електронске књиге у Србији

Данас велики број установа културе и образовања у Србији има своју дигиталну библиотеку или дигитални репозиторијум на интернету, који омогућава приступ књигама у електронском формату (IFLA/UNESCO, 2010). Ове базе података углавном настају дигитализацијом библиотечких збирки или постављањем електронских издања ових установа на мрежу и углавном се налазе у слободном приступу.

Први подухват покренут са циљем омогућавања приступа књигама у дигиталном формату у Србији био је Пројекат Растко, као српска верзија Пројекта Гутенберг. Пројекат Растко представља својеврсну онлајн библиотеку сачињену од књига из научних области и уметности које припадају српској и сродним културама, са циљем да се културна добра учине доступним домаћој и међународној јавности. На интернет презентацији пројекта наводи се да је овај подухват „непрофитан, невладин и волонтерски, а остварује се добровољном делатношћу појединаца и установа које се у ширем смислу баве проучавањем, развојем и очувањем српске културе“, ⁸ као и да је дигитализација књига извршена у складу са ауторским правима. Идеја о оснивању пројекта настала је 1994. године, а септембра 1997. пројекат је почео са јавним радом. Дигитална библиотека Растко данас садржи велики број књига у електронском формату,

⁷ Open Library, <https://openlibrary.org/>, приступљено 11. IV 2016.

⁸ Пројекат Растко: библиотека српске културе, <http://www.rastko.rs>, приступљено 22. III 2016.

у слободном приступу и са потпуно претраживим текстом. Што се ауторских права тиче, на сајту пројекта се наводи да је обим коришћења ауторских садржаја одређен дозволама носилаца ауторских права, што је дефинисано писаним споразумом, усклађеним са националном и међународном правном регулативом о ауторским правима. Временом је Пројекат Растко проширен и данас су у њега укључене и друге земље које дигитализују и постављају на мрежу електронске публикације које се односе на Србију и српску културу или се ње непосредно тичу. На тај начин читав пројекат је обogaћен и популаризован и ван граница Србије. Као део пројекта постоје: Пројекат Растко Темишвар: библиотека културе Срба у Румунији, Пројекат Растко Будимпешта-Сент Андреја: библиотека културе и традиције Срба у Мађарској, Пројекат Растко Кијев-Лавов: библиотека украјинско-српских културних веза, Пројекат Растко Бугарска: библиотека бугарско-српских културних веза и други.

Дигитална библиотека Филолошког факултета Универзитета у Београду⁹ настала је дигитализовањем издања Факултета, а тренутно је чини више од 107.000 страница дигитализованог материјала монографских и периодичних публикација. Библиотека је оформљена као подршка учењу на даљину и намењена је студентима, наставницима и истраживачима, а њен садржај је претражив по аутору, наслову, стручној и предметној класификацији и по свакој речи из текста.

Дигитална Народна библиотека Србије¹⁰ сачињена је од неколико збирки књига, часописа, фотодокумената и друге врсте грађе из фондова Народне библиотеке Србије, са циљем да омогући отворен приступ знању и информацијама. Део ове библиотеке је Српска дечија дигитална библиотека са 127 књига за децу, које су дигитализоване у склопу пројекта Међународна дечија дигитална библиотека.¹¹

Пројекат Антологија српске књижевности¹² има за циљ да свим заинтересованим читаоцима широм света учини лако доступним најзначајнија дела српске књижевности. Све књиге дигитализоване су поштујући ауторска права. Одабир књига начинио је Учитељски факултет Универзитета у Београду, а дигитализацију је извршио Мајкрософтов Центар за развој софтвера у Београду. Међу 130 бесплатно доступних књига могу се наћи дела Радоја

⁹ Дигитална библиотека, Филолошки факултет, <http://www.fil.bg.ac.rs/lang/sr/biblioteke/digitalna-biblioteka/>, приступљено 21. III 2016.

¹⁰ Дигитална Народна библиотека Србије, <http://www.digitalna.nb.rs/>, приступљено 21. III 2016.

¹¹ Међународна дечија дигитална библиотека, <http://en.childrenslibrary.org/>, приступљено 22. IV 2016.

¹² Антологија српске књижевности, <http://www.antologijasrpskeknjizevnosti.rs/>, приступљено 23. III 2016.

Домановића, Лазе Костића, Десанке Максимовић, Јована Стерије Поповића и других српских аутора. Захваљујући постојању електронске верзије ових књига, које се користе и као школска лектира, омогућен је бржи и истовремени приступ публикацији од стране већег броја корисника, чиме се решава проблем недовољног броја примерака књиге.

Приступ електронским садржајима који нуде неки од комерцијалних добављача у Србији је омогућен посредством Конзорцијума библиотека Србије за обједињену набавку (КоБСОН) који је оформљен 2002. године као једна од првих иницијатива за омогућавање приступа научним информацијама у електронском облику.¹³ КоБСОН је најпре био усмерен ка набавци електронских часописа, а касније је увео и могућност приступа књигама у електронском формату. Приступ електронским публикацијама преко КоБСОН-а има своја ограничења: може им се приступати са академске мреже, из институција чији је оснивач Влада Републике Србије и које су у надлежности Министарства културе, или, под посебним условима, од куће. У 2016. години посредством КоБСОН-а корисницима из Србије доступно је скоро 160.000 електронских књига из различитих научних области, од добављача Springer, EBSCO eBook, Google Books, Pub Med, ScienceDirect, Cleveland Medicine Index, DOAB и Hein on Line.

Иако је тржиште издавања електронских књига у Србији још увек недовољно развијено, ипак су направљени одређени помаци по том питању те данас код нас постоји неколико издавача који се баве издавањем књига у електронском формату. Један од првих издавача искључиво електронских књига у Србији била је издавачка кућа NM Libris, основана 2013. године. Функционисала је на следећи начин: део књига био је бесплатно доступан свима, а прављењем корисничког налога и плаћањем чланарине корисник је добијао могућност приступања и преузимања свих књига које су се налазиле на сајту. Поред тога, корисник је могао да направи своју библиотеку и да у њој организује књиге по својој жељи. Тренутно је сајт ове издавачке куће ван функције из непознатих разлога.

Српским издавачима електронских књига прикључила се и издавачка кућа Booka са два наслова: „У потпалубљу“ Владимира Арсенијевића и „39 песама“ Небојше Кривокуће, који су постављени на неколико сајтова специјализованих за продају електронских књига. Booka је најавила издавање и других књига у електронском формату.¹⁴ Издавачка кућа Архипелаг 15. маја 2014. године на своме сајту објавила је вест о покретању електронске издавачке делатности

¹³ Конзорцијум библиотека Србије за обједињену набавку, http://kobson.nb.rs/kobson/o_konzorcijumu.1.html, приступљено 21. III 2016.

¹⁴ Издавачко предузеће Booka, <http://www.booka.in/knjige/ebooks>, приступљено 21. III 2016.

у серији названој е-Архипелаг, уз навођење да је књига „текст који читамо и то је неупоредиво важније од самог облика којим се тај текст посредује до читалаца“.¹⁵ За сада је објављено осам електронских књига, а план је да се настави са објављивањем других наслова. Компанија Media Art Content DOO из Новог Сада такође се бави издавањем електронских књига у Србији.¹⁶ На своме сајту имају одређени број наслова, који се могу купити преко иностраних комерцијалних дистрибутера.

Иако, поред наведених, и други издавачи у Србији своју издавачку делатност проширују са штампаних на електронска издања, електронско издаваштво у Србији још увек није у потпуности заживело. О томе сведочи покушај претраге узајамног каталога Виртуелне библиотеке Србије у оквиру основног и изборног претраживања у потрази за подацима о књигама издатим у електронском облику. На овај начин није било могуће доћи до података који су нас занимали јер у сегментима основног и изборног претраживања у каталогу ВБС не постоји опција претраживања искључиво књига у електронском формату. Једина могућност је претрага у оквиру командног претраживања при чему је потребно направити посебан сложен упит. Претраживањем узајамног каталога упитом `EA="*/mon` добија се 10.648 публикација у електронској форми, међу којима и публикације на страним језицима, докторске дисертације и дигитализовани стари фонд. У покушају да дођемо до податка о електронској издавачкој делатности у Србији у протеклих 15 година, сузили смо претрагу и употребили упит (`EA="*/mon and LA=(scc or srp) and PY=20*) not TI=doktorsk*`). На овај начин добијено је 1690 резултата, односно толико монографских публикација у електронском облику издатих на српском језику од 2000. године до данас. Резултати претраге укључују, осим електронских књига, и каталоге, водиче и разна друга документа у електронској форми, а везе уз сваки добијени резултат воде до публикација које се налазе на мрежи. Оваквом претрагом могуће је донекле стећи увид у стање електронског издаваштва у Србији и пратити његов развој.

4. Уређаји за читање електронских књига

Да би могле да се користе, електронске књиге захтевају поседовање одговарајућег уређаја и софтвера. У раду „Електронске публикације и библиотеке – пресек стања“ говорећи о електронским књигама аутор наводи да

¹⁵ Издавачка кућа Архипелаг, <http://www.arhipelag.rs/arhipelag-magazin/gojko-bozovic-magazin-3/>, приступљено 22. III 2016.

¹⁶ Media Art Content LTD, <http://www.elektronske-knjige.org/>, приступљено 21. III 2016.

је „стално присутна и нека врста отпора према овој врсти публикација, будући да технологија још није на ступњу који би омогућио дуже и конформније читање“ (Васиљевић 2006). Са напретком технологије, ситуација се по том питању у међувремену променила. Електронске књиге могу се читати на рачунарима, лаптоповима, таблетима и паметним телефонима, а наведени уређаји сасвим добро могу послужити за краће читање. За дуже читање препоручују се уређаји који су примарно замишљени као електронски читачи књига. Са све развијенијим електронским читачима, читање електронских књига не разликује се много од читања штампаних издања. Ови уређаји омогућавају чување великог броја књига, које корисник у сваком тренутку може понети са собом и које може поделити са другима. Електронске књиге и читаче све чешће користе истраживачи и научни радници који су увидели предности оваквог начина читања и истраживачког рада (Nehring, 2013). Наиме, људима који добар део свога времена због посла проводе у путовању, практичније је да електронски читач напуне потребним књигама него да са собом носе неколико књига у штампаном издању.

У зависности од произвођача и модела, електронски читачи књига имају различите карактеристике. Постоје једноставни модели који омогућавају пребацивање књига са рачунара у одређеном формату, формирање личне библиотеке, претраживање текста и прављење белешки. Електронски читачи су путем мреже повезани са специјализованим е-продавницама у којима корисници могу куповати књиге, обично по нижој цени од штампане верзије. Поред оваквих једноставних, постоје и напреднији уређаји за читање који више имају карактеристике таблет рачунара и паметних телефона. Овакви уређаји садрже додатне апликације од којих неке нису намењене за читање (на пример, апликација за слушање радија или играње игрица). Пракса показује да овакви, напреднији уређаји, нису најбољи избор за читање књига у електронском формату, јер су они, фокусирајући се на додатне карактеристике, изгубили на квалитету када је реч о ономе што би примарно требало да буду – средства која ће омогућити што једноставније читање књига у електронској форми. Уређаји које су корисници оценили као најадекватније за дуже читање подржавају тзв. технологију електронског мастила. Ова технологија одаје утисак читања са папира, те се током читања не замарају очи као што се дешава код дужег читања са екрана монитора или таблета. Такође, овакав екран омогућава читање у мраку и на сунчевој светлости, без појављивања одсјаја и без треперења екрана. Предност технологије електронског мастила јесте боља издржљивост батерије, за разлику од уређаја који не користе ову технологију. Данас су на тржишту доступни уређаји са обичним екраном, као и они са екраном осетљивим на додир.

Данашњи електронски читачи појављују се у различитим облицима и са разним карактеристикама, па корисници могу да између разноврсних модела изаберу онај који највише одговара њиховим потребама. Што се тиче набавке уређаја за читање електронских књига у Србији, у неким продавницама се могу пронаћи одређени модели са различитим спецификацијама и по различитим ценама, али понуда је далеко од широке и разноврсне. Друга могућност набавке је онлајн куповина електронских читача из неке продавнице у иностранству. Као алтернатива набавци специјализованог уређаја за читање електронских књига, јавља се могућност читања са рачунара, таблета или паметног телефона. Данас постоји велики број апликација које омогућавају читање књига са наведених уређаја, који поред тога омогућавају манипулацију књигама и текстом унутар њих.¹⁷ Ове апликације донекле олакшавају иначе напорно читање са рачунара, али за најудобније читање књига у електронском формату свакако се препоручује електронски читач.

5. Предности и мане електронских књига и нови приступ читању

Научници и истраживачи желе брз и једноставан приступ информацијама а то им омогућавају управо електронске књиге. Нове генерације очекују да сву потребну литературу могу да пронађу онлајн (Adema, 2012). Двдесетчетворочасовна доступност једна је од предности електронских издања наспрам штампаних. Публикација у електронском облику која је постављена на интернет моментално постаје доступна свима. Неограничени број корисника може у исто време да чита електронску књигу, не морајући притом физички да буде присутан у библиотеци. Проблем недостатка довољног броја примерака публикације овим путем је решен.

Корисници се опредељују за читање електронских књига због тога што оваква издања не захтевају много простора за складиштење. Са појавом мобилних уређаја за читање, тзв. електронских читача, читање је добило нову димензију. Сада је могуће понети велики број књига свуда са собом. Поред тога, текст у електронском облику може се претраживати,¹⁸ могу се уређивати величина и фонт слова према личном укусу, а хиперлинкови могу да упућују са једне странице на другу (Трифуновић, 2005). У електронским књигама могу се правити белешке у тексту, а неки од електронских читача у себи садрже практичне електронске речнике.

¹⁷ Неке од таквих апликација су: Wattpad, Amazon Kindle, FBReader и Kobo eBooks.

¹⁸ Један од најзаступљенијих формата за кодирање електронских књига који омогућава да електронске књиге у пуном тексту буду у потпуности претраживе је међународни формат TEI (Text Encoding Initiative).

Електронске књиге могуће је брзо размењивати, копирати и пребацивати са једног уређаја на други. Као једна од предности електронских књига наводи се и њихова економичност и еколошка прихватљивост, јер њихово издавање не захтева потрошњу папира. Дигитализација штампаних издања омогућава презервацију и заштиту старе и ретке књиге и других публикација од честе и интензивне употребе. Осим тога, електронска издања старих и ретких књига неретко су једини начин да се широј јавности презентује национално и међународно културно наслеђе.

Мане које се приписују књигама у електронском формату више се односе на уређаје за њихово читање, него на саме публикације у електронском облику. Као једна од тих мана наводи се немогућност удобног читања током дужег временског периода. Читање књиге са рачунара отежавајуће је због реметилачких фактора као што су разне друге отворене апликације, прозори на рачунару и слично, који скрећу пажњу са самог процеса читања (Pettigrew, 2015). Преласком са монитора рачунара на електронске читаче књига, где су реметилачки фактори искључени или барем сведени на минимум, и ове замерке полако нестају.

У зависности од произвођача и модела уређаји за читање електронских књига подржавају само неке од формата па се то може сматрати ограничењем приликом коришћења електронских књига. Међутим, данас се на интернету могу пронаћи разноврсне апликације које читају различите формате електронских књига, те је стога ово ограничење такоређи занемарљиво. Програми као што су Calibre¹⁹ и Sumatra PDF²⁰ врло су једноставни за коришћење, а поред могућности пребацивања е-књига из формата у формат, корисницима омогућавају да креирају своју виртуелну библиотеку и да у њој манипулишу књигама. Батерија на читачима електронских књига може да траје прилично дуго и да издржи много сати читања.

Појава и коришћење електронских књига утицале су и на сам процес читања. Више аутора заступа став да се електронске књиге могу ефикасно користити у циљу промоције читања (Weber and Cavanaugh, 2006) и информационог описмењавања деце школског узраста (Schiemann, 2016). Књиге су данас, посредством интернета, доступније више него икад и сматра се да то може да допринесе „производњи“ читалаца који ће читати веома разнолика штива (Pettigrew, 2015) и повећавању свеукупне читалачке публике.

Првобитно замишљене као непосредна веза између аутора и читалаца, електронске књиге допринеле су развоју других врста интеракције. Иновације у области технологије допринеле су развоју тзв. интерактивних електронских

¹⁹ Calibre e-book management: <http://calibre-ebook.com/>, приступљено 22. III 2016.

²⁰ Sumatra PDF, <http://www.sumatrapdfreader.org/free-pdf-reader.html>, приступљено 22. III 2016.

књига, које омогућавају висок ниво интеракције између корисника, текста у електронском облику и дигиталног окружења (Bozkurt and Bozkaya, 2015). Поред креирања бележака у тексту електронски читачи нуде могућност брзог дељења забележеног садржаја путем интернета и друштвених мрежа, размену мисли и идеја и размишљања о прочитаном. На тај начин постиже се тренутна размена информација о прочитаном тексту, а читаоцима је, захваљујући праћењу садржаја који су креирали други, омогућено да на једноставнији начин изаберу адекватну литературу за читање и истраживачки рад (Софронијевић, 2015).

Књиге у електронском облику са својим карактеристикама допринеле су већој повезаности самог читаоца са текстом који чита. Аутори електронске књиге називају „прилагодљивим“, а као карактеристике њихове прилагодљивости наводе могућност подешавања величине текста, подвлачење, писање забелешки, интерактивне речнике и поседовање софтвера за читање наглас.²¹ Велика меморија електронског читача омогућила је читаоцу да своју библиотеку може увек и свуда да носи са собом уместо да буде ограничен на само неколико књига које може да има у сваком тренутку при руци. Захваљујући напредној технологији коју користе, уређаји за читање електронских књига практични су за свакодневно читање, јер се могу користити у различитим условима.

6. Закључак

С обзиром на то да живимо у времену када је технологија присутна у скоро сваком сегменту живота, њен утицај на активности као што је читање био је очекиван. Дигитално окружење допринело је промени у процесу читања тиме што је омогућило повезивање у мреже знања и информација, међусобну комуникацију читалаца и ближу везу са самим текстом који се чита. Све већи број електронских књига, прихватљива цена, лакши и једноставнији приступ и све напреднији електронски уређаји за читање, допринеће њиховој већој дистрибуцији и коришћењу. Поред предности као што су једноставан приступ, двадесетчетворочасовна доступност и економичност у смислу цене и простора, утицаће на то да се број људи који читају електронске књиге повећа, што ће допринети повећању укупне читалачке публике.

²¹ Софтвер за гласно читање електронских књига прилагођени су за читање на енглеском језику. Понеки софтвер такође омогућава читање текста на српском језику, али оно још увек није на задовољавајућем нивоу: речи се неразговорно изговарају и тешко се може пратити.

Литература

- Adema, Janneke. "On open books and fluid humanities". *Scholarly and Research Communication*, vol. 3, number 3 (2012):16, приступљено 12. IV 2016, <http://src-online.ca/src/index.php/src/article/view/92/107>.
- Bozkurt, Aras and Mujgan Bozkaya. "Evaluation criteria for interactive e-books for open and distance learning". *The International Review of Research in Open and Distributed Learning*, vol. 16, number 5 (2015):58–82, приступљено 24. IV 2016, https://www.academia.edu/16304364/Evaluation_Criteria_for_Interactive_E-Books_for_Open_and_Distance_Learning.
- Филипи-Матутиновић, Стела. *Ауторска права за библиотекаре: приручник*. Београд: Универзитетска библиотека „Светозар Марковић“, 2013.
- IFLA/UNESCO Manifesto for Digital Libraries, 2010. приступљено 13. IV 2016, <http://www.ifla.org/publications/iflaunesco-manifesto-for-digital-libraries>,
- Nehring, Daniel. "How Useful are E-readers for Academic Reading, Really?" приступљено 24. III 2016, <http://www.socialsciencespace.com/2013/03/how-useful-are-e-readers-for-academic-reading/>, 2013.
- Pettigrew, Sarah Ellen. "The future of the book and the slow book movement" in *History of the book: Disrupting Society from Tablet to Tablet*, 139–159. Western Oregon University, 2015, приступљено 23. III 2016, http://digitalcommons.wou.edu/cgi/viewcontent.cgi?article=1014&context=history_of_book.
- Софронијевић, Адам. „Нова парадигма сарадње у библиотекама“, Докт. дисертација, Универзитет у Београду, Филолошки факултет, 2015.
- Schiemann, Christopher. "E-books Can be a Great Choice for Middle Schoolers", in *School Library Journal*, February 2016, приступљено 14. IV 2016, <http://www.slj.com/2016/02/students/ebooks-can-be-a-great-choice-for-middle-schoolers/>.
- Трифуновић, Богдан. „Електронске књиге“. *Глас библиотеке*, бр. 12 (2005): 15, приступљено 21. III 2016, http://www.cacak-dis.rs/elektronska_izdanja/Glas_biblioteke_12/pdf_gb_12/Bogdan%20Trifunovic-Elektronske%20knjige.pdf.
- Васиљевић, Наташа. „Електронске публикације и библиотеке – пресек стања“. *Инфотека*, год. 7, бр. 1-2 (2006):75–80, приступљено 22. III 2016, http://infoteka.bg.ac.rs/pdf/Srp/2006/INFOTEKA_VII_1-2_Jun2006_75-80.pdf.
- Weber, Christine L and Terence W. Cavanaugh. "Promoting reading: Using ebooks with gifted and advanced readers". *Gifted Child Today*, year 4, vol. 29 (2006):56–63, приступљено 23. III 2016, <http://files.eric.ed.gov/fulltext/EJ746309.pdf>.

Вранеш, Александра и Марковић, Љиљана. *Од рукописа до библиотеке : појмовник*. Београд : Филолошки факултет, 2008.

Израда мултимедијалног документа „YU рок сцена“

УДК 004.55:378.147]:02(497.11)

САЖЕТАК: У овом раду ће бити представљен процес израде мултимедијалног документа под називом YU РОК СЦЕНА у коме су учествовали студенти завршне године основних академских студија на Катедри за библиотекарство и информатику Филолошког факултета Универзитета у Београду током школске 2014/2015. године у оквиру предмета Мултимедијални документи. У раду се приказује историјски развој рокенрола на просторима бивше Југославије, рок сцене у југословенским републикама, као и утицај рок музике на друге уметности и медије. Посебно се описује техничка реализација пројекта, као што је планирање, обрада, дизајн и креирање самог мултимедијалног документа.

КЉУЧНЕ РЕЧИ: мултимедијални документ, библиотекарство, информатика, рокенрол, музика, Југославија.

РАД ПРИМЉЕН: 28. март 2016.

РАД ПРИХВАЋЕН: 06. мај 2016.

Милена Обрадовић
milenaacm1@hotmail.com

Александра Арсенијевић
aki.7424@gmail.com

Михаило Шкорић
miccersoft@gmail.com

Универзитет у Београду
Филолошки факултет

1. Историјски развој рокенрола на просторима бивше Југославије

Рокенрол на просторима бивше Југославије вуче корене још из двадесетих година прошлог века када су почели да се појављују, тада нови, инструменти попут саксофона и гитаре (Fajfić i Nenad, 2006, стр. 18–21). Појава тих инструмената утиче да и југословенски музичари, у складу са светским трендовима, почињу да свирају џез. Ретки су били музичари који су до Другог светског рата свирали поменуте нове инструменте и џез ритмове, а током Другог светског рата долази до потпуног затишја. Тек после Другог светског рата почињу да се формирају ђачки и студентски оркестри који су редовно свирали џез на игранкама. Неки су у то време имали став, под утицајем медија и претрпљеног рата, да је џез „ђаволска музика“ која тера младе на бунт и да би је

требало забранити. Упркос оваквим ставовима цез је опстао, а томе је допринело и редовно слушање (цез) музике, укључујући и тадашње светске хитове, коју је емитовао *Радио Луксембург*. Па ипак, током педесетих година на сцени је више била присутна такозвана забавна музика.

Због сукоба са Совјетским савезом за време Хладног рата, Југославија, једна од земаља оснивача Покрета несврстаних, била је отворенија према Западу и свим производима поп културе, посебно америчке. Тако је Југославија била једина комунистичка земља која је учествовала на *Песми Евровизије* 1961. године. Почетком наредне деценије музичари, а нарочито омладина, почињу да прате, пре свега, британске и америчке трендове. Продавнице плоча биле су све чешће, а на одређеним местима музика се могла слушати преко дубокса. Тих година започета је и производња кока-коле на овим просторима, начин одевања се прилагођавао Западу, а појава цинса је изазвала праву револуцију у свету моде. Ни рокенрол није заостајао за овим трендовима, а тадашњи музичари су почели да праве обраде светских рок хитова, нарочито Елвиса Преслија, Чака Берија, Бадија Холија и многих других.

Тешко је рећи ко је био најзаслужнији за пробој рокенрола на југословенску музичку сцену. Многи сматрају да су Миле Лојпур и Перица Стојанчић из Београда и Карло Метикош из Загреба имали најзначајнију улогу у томе (Ђајић и Ненад, 2006, стр. 18–21). Рок се као музички правац веома брзо раширио по целој Југославији. Тада су се код нас развила два смера у рок музици: кантри поп рок (спој етно звука са рок стилем) и западни, најчешће англоамерички рок. За разлику од светске рок сцене чији су текстови били јако пропагандно средство против ратова и класних сукоба, првобитни домаћи рок текстови су имали визуелни и музички идентитет који се односио на младалачка маштања, снове о успеху, као и на генерацијске проблеме.

Број рок група и соло извођача је растао, а најплодоноснијим годинама се сматрају седамдесете године XX века, које су донеле потпуно омасовљавање и велику популарност рок сцени. Бендови настали почетком седамдесетих година XX века су били под великим утицајем хипи покрета, али од тада почиње гранање рок сцене на разне поджанрове, попут хард-рока, прогресивног рока, арт-рока, цеза, блуза и других. Крајем те деценије у Југославији се појављују панк (под утицајем британских и америчких панк бендова као што су *The Sex Pistols*, *The Clash*, *The Stooges*, *Ramones* и други) и нови талас као потпуно оригиналан културни и уметнички покрет.

Сам панк је антитеза устаљеним политичким, социјалним и економским принципима друштва, док је нови талас одисао посебним бунтом, променом свести и расположења југословенског друштва, које је постало отвореније и критичније, са слободнијим схватањима и укуси у једнопартијском политичком систему Југославије. Све до појаве панка и новог таласа главна

тема текстова је углавном била љубав, а са панком и новим таласом текстови су добили нову ширину и комплексност која се није уклапала у тадашње социјалне, културолошке и политичке норме. Мада су многи у то време нови талас поистовећивали са шундом (Kostić, 2013a), данас је изузетно цењен и сматра се да је то био феномен који се никад више није поновио.

Готово читаву наредну деценију је обележио нови талас, а пред крај осамдесетих година XX века су се појавиле групе које су биле под великим утицајем старијих колега. Неки извођачи су се кратко задржали на сцени, док су, са друге стране, многи мање или више, активни и данас. Свако од њих је оставио неизбрисив траг на југословенској музичкој сцени и имао велики утицај на друштво и културу у целини. Рок је у Југославији, као и у свету, представљао један свеобухватни покрет који је утицао на многобројна дешавања.

2. YU рок сцена по територијама

У овом делу Мултимедијалног пројекта је представљен велики број бендова и извођача са простора бивше Југославије, односно из Србије, Хрватске, Словеније, Босне и Херцеговине, Македоније и Црне Горе. У кратким биографијама група и извођача дате су за њих најважније године, чланови, манифестације на којима су учествовали, као и њихова комплетна дискографија.

2.1 YU рок сцена у Србији

У Србији центар оснивања и рада великог броја бендова, као и афирмисања соло извођача је био Београд. Током шездесетих година XX века на сцени се појављују састави као што су *Златни дечаки*, *Елипсе*, *Корни група*, *Породична мануфактура црног хлеба*, *Договор из 1804.*, *Искре*, *Црни бисери*, *дует Влада и Бајка*, *Центлмени*, *Пламених пет* и *Силуете*. На сцени се појављују и соло извођачи, од који су неки били и чланови неких од наведених група, попут Милета Лојпура, Перице Стојанчића, Душана Прелевића, Корнелија Ковача, Зорана Симјановића, Слободана Бобе Стефановића, Ђорђа Марјановића и многих других. Осим концерата и фестивала који се јављају у то време, попут фестивала *Београдско пролеће* или *Београдска гитаријада*, најважнија места за изласке су били клубови у којима се највише слушао рокенрол. Клубови као што су *Градски подрум*, *Дом омладине* и *Еуридика* била су омиљена места за окупљање омладине. Током седамдесетих година XX века се јављају бендови и извођачи на чију музику је посебно утицао хипи покрет. Ту се издвајају групе као што су: *YU група*, *Сунцокрет*, *Булевар*, *Последња игра лептира*, *Зана*, *С времена на време*, *Генерација пет*, *Рибља чорба*, *Игра стаклених перли*, *Лаки*

пингвини, али и соло извођача попут Лазара Ристовског, Срђана Марјановића, Оливера Мандића, Дејана Цукића, Слађане Милошевић, Александра Сање Илића и Рамба Амадеуса. Почетком осамдесетих на великој београдској рок сцени јављају се групе које стварају музику под утицајем пре свега новог таласа. У те групе спадају *Идоли*, *Шарло акробата*, *Електрични оргазам*, *Екатарина Велика*, *Дисциплина кичме*, *Партибрејкर्स* и *У шкрипцу*. Тада се такође јављају и други рок састави: *Пилоти*, *Викторија*, *Цакарта*, *Руџ*, *Дивљи анђели*, *Алиса*, *Октобар 1864*, *Сирова кожа*, *Бајага и Инструктори*, *С.Т.Р.А.Х.*, *Ван Гог* и *Вампири*. Од соло извођача посебно се издвајају Владимир Влада Дивљан, Беби Дол, Милан Делчић Делча и Тони Монтано.

Иако је у Србији Београд био епицентар културних дешавања и креирања рок музике, треба истаћи да ни остатак Србије није заостајао. Већи градови у Србији значајни за рок културу Југославије, који су изродили велике рок бендове, су: Ниш, Чачак, Крагујевац и Нови Сад, као и мањи градови, попут Бечеја (одакле потиче група *Ева Браун*), који су, са по једним или два бенда, ушли у рок историју бивше Југославије. У Нишу су настале значајне групе – *Галија*, *Кербер* и *Добри Исак*, а за време Југославије, тамо су се одржавали велики рок концерти и свирке. У Чачку су у то време основане 84 рокенрол групе и већина њих је настала из Културно-уметничког друштва *Абрашевић*. Први и најпопуларнији чачански састав шездесетих био је Беле вишње, а након њих се појављују нове групе: *Звечарке*, *Црне мамбе*, *Сребрне сенке*, *Дечац са Мораве*, *Хермелини*, *Чачански племићи*, *Сафири*, *Наутилус*, *Кућни савет*, *Себастијан*, *Морис*, *Мрачни Мо*, *Спин*, *Високи напон*, *Крвна група* и друге. Из Чачка воде порекло два врхунска рок мајстора: Радомир Михаиловић Точак и Борисав Ђорђевић – Бора Чорба. Крагујевац је временом постао синоним за групу *Смак*, један од најстаријих и најзначајнијих рок бендова са ових просторима, а из Крагујевца су такође и састав *Освајачи*, као и панк група *КБО!*. Један од најбитнијих представника Новог Сада је Ђорђе Балашевић, који је, пре почетка своје успешне соло каријере, био члан група *Жетва* и *Рани траз*. Из Новог Сада је и група *Гарави сокак*.

2.2 YU рок сцена у Хрватској

Први талас рокенрола у Хрватској је кренуо крајем педесетих, међутим, први прави рок састави, нарочито из Загреба, Пуле, Ријеке и Сплита, оснивају се почетком шездесетих година, а међу најпознатијима су: *Бијеле стријеле*, *Група 220*, *Црвени корали*, *Делфини*, *Безимени*, *Атоми*, *Квартет 4М* и многи други. Један од првих великих и светски познатих хрватских соло извођача, који се сматра и посебно заслужним за представљање и ширење рок музике на овим просторима, јесте Карло Метикош, познатији као Мет Колинс. Осим

његу, тада успешне соло каријере почињу да граде и Ивица Перџл, Јосипа Лисац, Зденка Ковачичек, Драго Млинарец и др. Најважнији бендови у Хрватској седамдесетих година, од којих су неки припадници и новог таласа су *Азра*, *Хаустор*, *Филм*, *Парни ваљак* и *Прљава казалиште*, а уз њих су активни и *Атомско склониште*, *Аеродром*, *Тајм*, *Параф*, *Други начин*, *Метак* и многе друге. Најважнији панк-рок бендови тог периода у Хрватској су *Параф* и *КУД Идијоти*. Неки од чланова ових бендова касније настављају веома успешну соло каријеру, као на пример Дато Топић, Дино Дворник, Бранимир Џони Штулић, Дарко Рундек и Масимо Савић. На хрватској рок сцени се осамдесетих појављују, између осталих и *Психомодо поп*, *Неки то воле вруће*, *Доријан Греј*, *Осми путник*, *Ђаволи* и *Фит*. Крајем осамдесетих и почетком деведесетих хрватска музичка сцена постаје права индустрија, која производи веома популарне бендове.

2.3 YU рок сцена у Словенији

У Словенији су увек за слушање биле најпопуларније шансоне и фолк музика, међутим, шездесетих година XX века, као и било где другде на простору бивше Југославије, у Љубљани, Копру, Марибору и Крању, полако почиње да се појављује, под утицајем Запада, све већи број рок састава, и тада настаје права културна рок револуција. Почињу да се праве посебне емисије на радију и пишу се чланци посвећени овом новом музичком правцу. Један од првих словеначких рок бендова који се пробио на велику рок сцену читаве Југославије је био састав Камелеони. Рок музика у Словенији постаје преовлађујућа на субкултурним догађајима попут фестивала, свирки и других окупљања. Међу познатијим словеначким рок и панк бендовима су: *Беле вроне*, *Булдожер*, *Лачни Франци*, *Лајбах*, *Панкрти*, *Видеосекс* и многи други, а међу значајније соло извођаче спадају и Томаж Домицељ, Марјета Неца Фалк, Андреј Шифрер и други.

2.4 YU рок сцена у Босни и Херцеговини

Босна и Херцеговина је у Југославији била плодно тло за развој поп и рок музике, што се посебно видело у Сарајеву, Травнику, Бугојну и Бихаћу. Један од првих рок састава на овим просторима били су *Индекси*, којима се касније прикључио и Корнелије Ковач. Након *Индекса*, на босанскохерцеговачкој рок сцени се појављују бројне популарне групе, као што су: *Про арте*, *Тешка индустрија*, *Амбасадори*, *Кодекси*, *Бијело дугме*, *Формула 4*, *Конгрес*, *Дивље јагоде*, *Ватрени пољубац*, *Забрањено пушење*, *Бомбај штампа*, *Елвис Ц. Куртовић* и његови *Метеори*, *СЦХ*, *Валентино*, *Плави оркестар*, *Мерлин*, *Хари Мата Хари*, *Црвена јабука* и многи други, као и соло извођачи који су пре или

касније били чланови поменутих група: Сеид Мемиф Вајта, Здравко Чолић, Исмета Дервоз-Крвавац, Јадранка Стојаковић, Харис Варешановић, Жељко Бебек, Горан Бреговић, Младен Војичић Тифа, Ален Исламовић, Елвис Џ. Куртовић, Дино Мерлин и други. Посебно треба издвојити Желимира Алтарца Чичка, једног од најзначајнијих и најутицајнијих креатора сарајевске музичке сцене, који је као уредник и водитељ бројних популарних радио и ТВ емисија, али и као организатор многих музичких манифестација крчио пут до успеха бројним неафирмисаним рок бендовима. Многи од поменутих група и извођача постали су веома цењени у свету. Иако је распад Југославије највише утицао на будућност група из Босне и Херцеговине, музику коју су стварали и данас радо слушају све генерације са ових простора.

2.5 YU рок сцена у Македонији

У Македонији је, нарочито у Скопљу, рок музика почела да се утемељује шездесетих година прошлог века. Међу првим саставима појављују се *Безимени* и *Бисери*, који се потом удружују у састав *Бис-Без* и почињу да уводе примесе етно звука. Њихова оригиналност је касније имала утицаја на композиције најпознатијег македонског бенда *Леб и сол* на чијем челу се налази светски познат и изузетно цењен музичар Влатко Стефановски. Поред наведених, на македонској рок сцени су је јавили још многи састави, а међу познатијима су: *Александар Македонски*, *Анастасија*, *Архангел*, *Бастаион*, *Ден за ден*, *Мизар* и други. Значајан допринос ширењу рок музике у Македонији и пробијању младих и неафирмисаних рок бендова дао је Радио Скопје, који је емитовао како музику са простора бивше Југославије, тако и тадашње светски познате рок хитове.

2.6 YU рок сцена у Црној Гори

Рок музика је од свих земаља бивше Југославије понајмање заживела у Црној Гори. Разлози за то су бројни, али најзначајнији су културолошки и материјални аспекти, као и чињеница да је тада у Црној Гори било тешко набавити добре музичке инструменте по приступачној цени. Ипак, постојале су тенденције да се рок музика укорени као музички правац на простору Црне Горе, посебно у Подгорици. О томе говоре неки филмови, попут филма Момира Матовића *То лудо срце*, као и рок састави *Ентузијастички*, *Стеле*, *Џефердари*, *Лордови*, *Свјетла у тами*, *Мртве душе*, *Страшине сјенке*, *Подгоричани*, *Ноћни путници* и други. Ови бендови углавном нису заживели, јер су само неки од њих успели да сниме по који демо снимак и данас се мало шта о њима зна. Тако они нису успели да се пробију ни на сопственој, а камоли на некој већој музичкој сцени. Највећи рок музичар Црне Горе је Миладин Шобић који се активно бавио

музиком почетком осамдесетих година XX века. Иако нису имали сопствену рок сцену, млади су били под великим утицајем рок музике са осталих простора бивше Југославије.

3. Утицај рокенрола на уметност, медије и друштвене догађаје

Утицај рокенрола се веома брзо осетио у многим друштвеним сферама. Заступљеност рок музике у медијима утицала је на њену све већу популарност. Преко радио и ТВ емисија, часописа, књига и филмова, рокенрол је допирао до широког круга људи. Убрзо почињу да се организују фестивали посвећени рок музици, као и да се одржавају свирке у све већем броју дискотека широм Југославије. Оснивају се прве дискографске куће, а рок музику почињу да прате и бројне контроверзне појаве. У овом раду одлучили смо да представимо најважније аспекте утицаја рокенрола на медије и друштвене догађаје, тако што смо их поделили на следеће теме, које уједно одговарају и подели у самом пројекту: Свирке и дискотеке; Фестивали; Поезија и књиге; Албуми и дискографске куће; Култне радио емисије; Култне ТВ емисије; Филмографија; Контроверзе и Часописи.

3.1 Свирке и дискотеке

Од 1977. године у Студентском културном центру (СКЦ) појављују се бендови као што су *Лимуново дрво*, *Игра стаклених перли*, *Електрични оргазам*, и на тај начин полако се формира полазна тачка експерименталног позоришта и алтернативног рока – „новог таласа“. Како се поменути „нови талас“ ширио, у СКЦ-у је наступало све више бендова, међу којима се издвајају *Панкрти*, *Шарло акробата*, *Идоли*, *Урбана герила*. Уследио је период комерцијализације, наплаћивања улазница, па и отварање све већег броја дискотека широм Југославије.

У самом Београду је постојало више дискотека. У *Цепелин* су претежно излазили млади људи, којима је хипи био близак начин размишљања, док су се у *Акваријусу* окупљали нешто озбиљнији припадници рок популације.

3.2 Фестивали

Рокенрол се као правац развијао од скромних почетака младих бендова Југославије који су учествовали на првим свиркама и игранкама, до првих концерата и фестивала. Период рок фестивала у Југославији почео је 1961.

године са Омладинским фестивалом у Суботици. Многи фестивали имали су кампање са лецима и постерима. Рок фестивали су имали за циљ, пре свега, спајање људи, промовисање рока и ширење његовог утицаја. Неки од најзначајнијих југословенских рок фестивала који су се у Југославији одржавали су: *Парада ритма*, *Ватромет ритма*, *I фестивал Beat музике*, *II фестивал ВИС-ова*, *II фестивал ВИС-ова*, *I шампионат ВИС-ова Југославије*, *I београдска гитаријада*, *II београдска гитаријада*, *I сарајевска гитаријада*, *Такмичење београдских и сарајевских ВИС-ова*, *Југословенски фестивал Beat музике*, *I фестивал југословенске поп музике*, *II фестивал југословенске поп музике*, *Воот фестивал*, *Гитаријада у Зајечару*, *Омладински фестивал у Суботици* и *ФАМУС* (Фестивал акустичарске музике у Сивцу).

3.3 Поезија и књиге

Рок музика пронашла је своје место и у књигама и поезији. У пројекту Мултимедијални документ 2014/2015 наведено је и укратко представљено око 70 наслова публикација које се баве рок сценом Југославије. Публикације попут *Ilustrovane YU rock enciklopedije 1960-2006* Петра Јањатовића (Јањатовић, 2001) и *Leksikona YU mitologije* (Andrić i dr., 2012) описују развој рок музике у Југославији, дају биографије и дискографије извођача, радио и ТВ емисије, фестивале и сл. Треба издвојити и књиге *Љубав је само реч: избор љубавне поезије југословенске забавне и рок музике* приређивача Војислава Несторовића (Несторовић, 2003) и *Обична љубавна песма* приређивача Ненада Атанасковића и Мирослава Јосиповића (Atanasković i Josipović, 2004). Књиге које су наведене и описане у овој посебној теми, биле су нам од велике помоћи приликом израде целокупног пројекта. Захваљујући раду аутора представљених публикација, доступне су нам многобројне информације о свим аспектима југословенске рок сцене.

3.4 Албуми и дискографске куће

Прве дискографске куће у Југославији појављују се после Другог светског рата. Југотон је прва издавачка кућа у Југославији, основана 1947. године. Производња грамофонских плоча у Југославији је започела 1956. године.

У Београду је 1951. године основана издавачка кућа ПГП РТБ. Међу најважније албуме које је ова издавачка кућа издала, спадају *Пуб Ђорђа Балашевића* (1982) и *Ујед за душу Рибље чорбе* (1987).

Издавачка кућа *Дискотон* основана је у Сарајеву 1973. године. За њих су албуме издавали *Бијело дугме*, *Забрањено пушење*, *Индекси*, *Дивље јагоде* и многи други.

У Загребу је 1972. године са радом започела агенција Suzy, која није имала сопствени погон за израду плоча, али је користила услуге других произвођача. Агенција је, између осталог, организовала концерте *Парног ваљка*, *Азре*, *Прљавог казалишта* и других познатих извођача.

У књизи *YU 100: najbolji albumi jugoslovenske rock i pop muzike* Душка Антонића и Данила Штрпца (1998) (Kostić, 2013b), као музичке групе са највише албума на листи издавајају се *Бијело дугме* и *Рибља чорба*, обе са по 8 албума. Издавачка кућа Југотон издала је највише албума, чак 47.

3.5 Часописи

Часописи са рок тематиком брзо су у Југославији стекли велику популарност, посебно код млађе популације. Извештавали су о концертима, доносили постере и топ листе најслушанијих песама, као и веома садржајне интервјуе са извођачима. Неке од највећих звезда о којима су часописи писали су *Рибља чорба*, *Индекси*, *Идоли*, а од страних звезда најзаступљенији су били Џон Ленон, Боб Марли и многи други. Неки од часописа који су излазили у Југославији су: *Ритам*, *Дубокс*, *Ладин џубокс*, *Тао*, *Бум*, *Време забаве*, *Ју рок магазин*, *Укус несташних*, *X zabava*, *Pop Express*, *Хероина* и *Рок 82*.

3.6 Култне радио емисије

“Нови талас” се из Београда и Загреба проширио по целој Југославији. Присутност рок музике на радију била је од кључног значаја за пробој извођача и представљање њихових хитова што већем броју људи. Личности које су дале велики допринос радио емисијама везаним за YU рок сцену су Зоран Модли, Никола Караклајић, Никола Нешковић, Влада Цет, Никола Кањевац и многи други. Радио Београд је у то време емитовао популарне емисије, као што су: *Зелени мегахери*, *Вече уз радио*, *Пријатељи звезда*, *Састанак у 9 и 5*, *То је само рокенрол*. Двестадвојка је емитовала *Вентилатор* и *Хит 202*, а Студио Б *Дискомер*. Радио Нови Сад емитовао је *Рандеву с музиком*, а Радио Пожаревац емисију *То је само поп и рок*. Изузетно популарна радио станица широм Европе, па и у Југославији, био је *Радио Луксембург*.

3.7 Култне ТВ емисије

Рок музика је после радија и часописа, своје место добила и на телевизији. Када је Радио Београд прерастао у Радио-телевизију Београд 1958. године, нову телевизију су у Србији, као и у целој у Југославији, претежно формирале професије са радија, филма, позоришта и штампе. Програми су пратили новине

у телевизијском стваралаштву, као и развој технологије. Приказиване су игране и драмске серије, документарни, забавни, образовни програми, програми за децу и филмови. Када је рок музика у питању, после повремених осврта у ТВ Дневнику, покренута је и прва емисија која се бавила рок сценом – *Концерт за луди млади свет*. Широј публици је на тај начин било омогућено да се информише о концертима, новим хитовима и албумима све већег броја извођача. Међу најважније култне ТВ емисије, које су редовно или повремено извештавале о свему везаном за југословенску рок сцену, спадају *Хит месеца*, *Стереовизија*, *Рокенролер*, *Формула 1*, *Београд ноћу*, *Петком у 22*, *Концерт за луди млади свет*, *Rock око*, *Топ листа надреалиста*, *Максиметар* и *Гаража*.

3.8 Филмографија

Почевши од касних педесетих, па све до средине шездесетих година XX века, музика се ненаметљиво провлачила кроз филмове *Суботом увече* (1957) Владимира Поповића, *Љубав и мода* (1960) Љубомира Радичевића и *Звиздук у осам* (1962) Саве Мрмка. У филмовима *Кад будем мртав и бео* (1967) Живојина Павловића и *Немирни* (1969) Кокана Ракоњца, захваљујући групама као што су *Елипсе*, *Златни дечаки* и *Силуете*, рок музика бива све више заступљена. Осамдесетих година прошлог века, у филмовима *Давитељ против давителја* (1984) Слободана Шијана, *Дечко који обећава* (1980) Милоша Радивојевића, *Црна Марија* (1986) Милана Живковића и *Шест дана јуна* (1985) Динка Туцаковића, спој рока и филма достиже врхунац. Међу музичарима који су допринели сједињењу рок музике и филма, посебно се истичу Корнелије Ковач, Војкан Борисављевић, Зоран Симјановић и Влатко Стефановски. Многи музичари су глумили у филмовима, а глумци су постајали музичари.

3.9 Контроверзе

Неки музичари су проналазили начин да изразе своје мишљење и ставове, као и да се додатно истакну на музичкој сцени, користећи се за оно време контроверзним текстовима песама, омотима албума, одећом или шминком. Тако је јавност, између осталог, била уздрмана песмама као што су *Курвини синови* (Азра), *Црни мерцедес*, *На западу ништа ново* (Рибља чорба) и *Краљ алкохол* (група *Time*). Велику пажњу изазвао је и омот албума *Рибље чорбе* – *Покварена машта и прљаве страсти*. Дуга коса, необична шминка, текстови песама који отворено указују на проблеме у друштву или на теме које су биле табу, све је то на неки начин сматрано скандалозним.

4. Прикупљање, обрада и обликовање садржаја

Од грађе за овај мултимедијални документ прикупљени су, пре свега, многобројни текстови из енциклопедија, хемеротека, као и примарних и секундарних монографских извора. Остали текстови, фотографије, видео и звучни записи прикупљени су делом из институција медија и културе (библиотека, хемеротека, архива, радио станица и телевизијских кућа, атељеа и кинотека), а делом са веба. Током периода од четири месеца посећено је преко двадесет институција у седам градова, и успостављена је сарадња са многобројним појединцима коју су помогли овом пројекту. Истраживање и прикупљање грађе обавили су студенти подељени у групе.

Један део група је истраживао територију или делове територија бивших југословенских република: Београд (Александра Арсенијевић, Милица Нинковић и Милена Обрадовић), остатак Србије (Александра Којић), Хрватска (Јована Дошенивић, Маја Иванчић и Марко Петровић), Босна и Херцеговина (Виолета Колаковић, Милица Перишић и Петар Поповић), Словенија (Михаило Шкорић), Македонија и Црна Гора (Александар Павловић).

Преостале групе су се бавиле утицајем рокенрола на уметност, медије и друштвене догађаје: свирке и дискотеке (Катарина Југов, Ивана Мердовић и Јелена Радић), рок фестивали (Милан Авдулах и Урош Тошић), поезија и књиге (Марија Брашанац и Тамара Станојчић), албуми и дискографске куће (Сњежана Мирчета и Иван Пеић), часописи (Маја Миладиновић и Катарина Радовановић), култне радио емисије (Милица Ђорђевић), култне ТВ емисије (Ивана Мердовић и Милена Обрадовић), филмографија (Александра Арсенијевић и Јелена Радић) и контроверзе (Владана Нешовић и Урош Тимић).

После завршеног прикупљања, материјал је требало на неки начин обрадити, систематизовати и на жељени начин приказати. Овим послом су се бавили Михаило Шкорић и Петар Поповић.

Најпре је почела обрада материјала који се односио на регионалну поделу сцене и изабране рок групе које су из тих региона потекле. Чланови група студената које су се овим темама бавиле, добили су задатак да своје прикупљене материјале организују на задат начин како би сав материјал могао да се обради на истоветан начин. У ту сврху је коришћена децимална класификација садржаја – сваком конкретном елементу материјала је придружен број као јединствени идентификатор (одредница). Прва цифра идентификатора се односи на регион, друга на град из тог региона, трећа и четврта на бенд из тог града, а пета и шеста на албум тог бенда. Подаци о материјалу и придружене одреднице су најпре унети у табеле програма Excel, а затим трансформисани у SQL базу података.

Табела 1 илуструје на који начин се у *SQL* бази података чувају подаци о рок групама. Колона *GID* (идентификатор групе) је придружени идентификатор групе (прва јединица је Србија, друга Београд, а трећа и четврта – редни број приликом набрајања група), док остале колоне садрже назив групе (истоимена колона), *MID* (идентификатор места – прва јединица је Србија, друга Београд), као и називе датотека које редом представљају слику чланова групе или њиховог албума, биографију, дискографију и одломак из једне музичке нумере одговарајуће групе. Редундантност података у овој бази је очигледна (идентификатор се понавља у називима датотека), али је она била неопходна ради повећања контроле и лакшег проналажења и исправљања многобројних грешака као последица чињенице да је базу попуњавало више лица. После усклађивања садржаја базе података са пратећим датотекама, сав материјал и база података су пребачени на локални сервер ради развијања и тестирања веб-сајта за преглед и претрагу сакупљеног материјала.

GID	Назив групе	MID	Слика	Биографија	Дискографија	Нумера
1101	Дисциплина кичме	11	1101.jpg	1101b.txt	1101d.txt	1101.mp3
1102	Дивљи анђели	11	1102.jpg	1102b.txt	1102d.txt	1102.mp3
1103	Договор из 1804.	11	1103.jpg	1103b.txt	1103d.txt	1103.mp3

Табела 1. Пример дела табеле *grupe*

Слика 1. Пример странице за регион – Словенија

Најпре је састављена основна композиција веб сајта у чијој је сржи била хијерархијска архитектура. Уместо да се за сваки регион, град, бенд праве

засебне *HTML* странице, направљене су четири *PHP*-скрипте које консултују *SQL*-базу са метаподацима о прикупљеном материјалу и, у зависности од избора корисника, динамички генеришу одговарајући *HTML*-приказ.

Прва скрипта, *tara.php*, приказује мапу бивше СФРЈ са републичким границама и главним градовима. Ознака главног града је уједно и веза ка скрипти *rep.php* са одговарајућим параметрима.

Скрипта *rep.php* описује рок сцену једног конкретног региона (републике), генерише називе свих градова тог региона и њихове слике. Слика сваког града представља везу ка опису рок сцене у одговарајућем граду (Слика 1). Испод листе градова *rep.php* исписује текст о рок сцени у изабраном региону, док се изнад листе приказују везе ка осталим регионима.

Активирањем везе на одређени град се корисник преусмерава на скрипту *grad.php*, која генерише листу извођача тог града са пратећим сликама. На врху генерисане странице се налази навигација (везе ка осталим градовима региона у коме се налази изабрани град) (Слика 2). У позадини се приказује слика изабраног града. Свака слика извођача је веза ка страници *bend.php* са одговарајућим параметрима (Слика 3).

Последњи скрипт, *bend.php* садржи само једну везу – за повратак назад. Улога ове скрипте је да у два одвојена одељка прикаже биографију и дискографију изабраног извођача или бенда. За сваки албум одабраног извођача се наводи назив, годиште и издавач, а уколико постоји, приказана и је слика омота одговарајућег албума. По учитавању странице, аутоматски се покреће репродукција унапред одређене музичке нумере одговарајућег извођача, а у позадини странице се налази његова слика.



Слика 2. Пример странице за град – Љубљана



Слика 3. Пример странице за извођаче – Булдожер

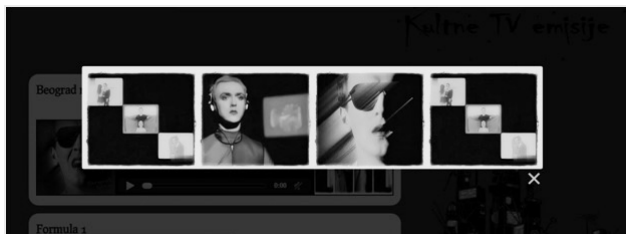
Како постоји велики број могућих исхода страница, неопходно је било пронаћи им одговарајућу заједничку тему (боју). За доминантну боју изабрана је нијанса црвене (\#dd2233), уз серифни фонт *prociono*. За боју заглавља је одабран дезен јагоде. Заглавље је заједничко за све странице мултимедијалног документа и састоји се из главног менија, контроле за репродукцију звука (audio player) и контрола за претрагу. Ови саставни делови остављени су да плутају применом CSS-својства *float*, чиме се жртвује флексибилност зарад изгледа, под претпоставком да ће документ бити прегледан на рачунару стандардне резолуције.



Слика 4. Изглед почетне странице

Под истом претпоставком је направљена и уводна страница као приказ собе (Слика 4) која, осим мапе за избор региона, садржи друге предмете као везе које упућују корисника на странице са описима специјализованих области које чине музичку сцену поред самих извођача: фестивале, свирке, дискотеке,

музичке часописе, филмографију, радио и тв емисије, дискографске куће, књиге и контроверзе које је југословенска рок музичка сцена проузроковала.



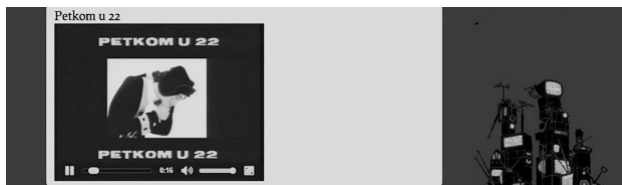
Слика 5. Пример приказа галерије са корисничке стране

Странице са описима специјализованих области садрже текстове које су студенти написали, као и везе на различите мултимедијалне садржаје: слике, аудио и видео записе. Позадина ових страница обојена је у нијансу зелене боје (\#E1F5A9), често праћена одговарајућом позадинском сликом. Мултимедијални материјали су представљени помоћу *Javascript* програма који се покреће кликом на жељену слику или запис. Све слике из галерије се могу листати, а новоотворени приказ се може угасити кликом на крстић (x) (Слика 5). Видео снимци у мултимедијалним галеријама се покрећу у видео плејеру, а звучни у аудио плејеру који су уграђени у прелистач (Слика 6).

Као што је већ поменуто, заглавље сваке странице садржи контроле за претраживање метаподатака о материјалу (називе извођача, албума и места). Уколико корисник не зна одакле је неки извођач или не жели да пролази кроз цео документ, везу ка биографији извођача може добити уколико укуца део назива у поље за претрагу. На основу корисничког захтева се генерише упит над базом података, како би се пронашли називи република, градова, извођача, албума који у свом имену садрже задати текст (Слика 7).

5. Студентско виђење пројекта – добре и лоше стране

Рад на пројекту под називом *YU РОК СЦЕНА*, у оквиру предмета Мултимедијални документи 2014/2015, био је за све студенте завршне године библиотекарства и информатике искуство од непроцењивог значаја. Били смо у прилици да објединимо и применимо знања стечена током четворогодишњих студија, али и да стекнемо нова. Тема пројекта је веома занимљива и нама блиска. Успешно смо сарађивали са значајним институцијама и појединцима.



Слика 6. Пример приказа видео записа

Пројекат Мултимедијални документ 2014/2015. омогућио је сваком од нас да искаже своја интересовања и способности.

Са друге стране, отежавајућа околност била је обимност теме. Избор и организација информација представљали су нам велики изазов. Трудили смо се да одаберемо најважније информације и да их представимо на разумљив начин. Будући да је на овом пројекту радило много студената, неопходна је била добра организација и смисао за тимски рад.

Мултимедијални документ је доступан на сајту Филолошког факултета на адреси http://www.fil.bg.ac.rs/mmd_27/mmd_2015/home.php.

```
<?php
$query = "SELECT Naziv, Ime, grupe.GID, Godina, Omot FROM albumi, grupe
WHERE albumi.GID = grupe.GID AND INSTR(Naziv, '{$grad}') > 0";
$result = mysql_query($query, $link);
if($result === FALSE)
{
    die("Upit nije vratio ništa");
};
$count = 0;
while ($row = mysql_fetch_array($result, MYSQL_BOTH))
{
    if($count==3)
    {
        echo "<tr>";
        $count = 0;
    }
    if($count==0)
    {
        echo "<tr>";
        echo "<td>";
    }
    echo "<p class='dgra'>" . $row['Naziv'] . " , </p></a>";
    echo "<p class='dgra'>" . $row['Godina'] . " </p></a>";
    echo "<a href='bend.php?x=" . $row['GID'] . "' class='dgra'><p>" . $row['Ime'] . " </p></a>";
    $adresa = "CMOTI/" . $row["Omot"];?>
```

Слика 7. Скрипта која претражује и исписује пронађене извођаче

6. Захвалност

Велику захвалност за израду овог пројекта дугујемо професорки Цветани Крстев, која је носилац пројекта, као и професору Милошу Утвићу и координаторки Биљани Лазић. Њихова помоћ, усмеравања и идеје били су

нам од непроцењивог значаја. Захваљујемо се и на подршци и сарадњи појединцима и институцијама који су нам уступили на коришћење богату грађу и поделили са нама значајне информације о југословенској рок сцени: Петру Јањатовићу, Жељку Радишићу, Мирољубу Стојановићу, Ивану Велисављевићу, Бориславу Станојевићу, Марку Маршићевићу, Нади Петронијевић-Човић, Јовану Јовановићу, Љубинки Гавран, Еду Грацину, Бранку Вурушићу, Радоману Кањевцу, Зорану Модлију, Срђану Николићу, Слоби Коњовићу, Љиљани Крнети, Војину Шарчевићу, Дарку Копјану, Марку Савићу, Јелени Стевановић и Слободану Боби Новаковићу, као и институцијама: Радио-телевизији Србије, Филмском центру Србије, Народној библиотеци Србије, Библиотеци града Београда, Радио Нови Сад, Радио Београд, Радио 202, Радио Пожаревац, Хрватској радиотелевизији, Југословенској кинотеци, Националној и свеучилишној књижници у Загребу, Народној и универзитетској библиотеци Словеније (Narodna in univerzitetna knjižnica) и Студентском културном центру.

Литература

- Andrić, Iris, Vladimir Arsenijević, i Đorđe Matić, ur. *Leksikon YU mitologije*. Beograd : Rende, Zagreb : Postscriptum, 2012.
- Atanasković, Nenad i Miroslav Josipović, ur. *Obična ljubavna pesma : antologija YU Rock ljubavne poezije*. Beograd : Mono & Mañana, 2004.
- Fajfić, Željko i Milan Nenad. *Istorija Yu rock muzike*. Lačarak : Tabernakl, 2009.
- Janjatović, Petar. *Ilustrovana YU rock enciklopedija : 1960–2000..* Beograd : Pro-metej, 2001.
- Kostić, Petar. „Šta su nama kič i šund?”. <http://balkanrock.com/>, линк sta-su-nama-kic-i-sund, приступљено 20. III 2016, 2013a.
- Kostić, Petar. „Yubilej: YU 100 – najbolji albumi jugoslovenske rok i pop muzike (15 godina kasnije)”. <http://balkanrock.com/>, линк yubilej-yu-100-najbolji-albumi-jugoslovenske-rok-i-pop-muzike-15-godina-kasnije/, приступљено 20. III 2016, 2013b.
- Несторовић, Војислав. *Љубав је само реч: избор љубавне поезије југословенске забавне и рок музике*. Београд : Београдска књига, 2003.

Рачунарство у друштвеним наукама: приказ студијског програма мастер студија

Владан Деведић

devedzic@gmail.com

Универзитет у Београду

Факултет организационих
наука

Милан Крстић

milan.krstic.fpn@gmail.com

Универзитет у Београду

Факултет политичких
наука

РАД ПРИМЉЕН: 03. септембар 2015.

РАД ПРИХВАЋЕН: 05. септембар 2015.

Студијски програм мастер академских студија „Рачунарство у друштвеним наукама“ на мастер студијама Универзитета у Београду развијен је као резултат Темпус пројекта INCOMING (Interdisciplinary Curricula in Computing to Meet Labor Market Needs, број пројекта 530155-TEMPUS-1-2012-1-EE-TEMPUS-JPCR). Осмишљен је као наставак основних академских студија у некој од области друштвених наука (економске науке, социологија, психологија, правне науке, финансије, менаџмент, филолошке науке, итд. или као комбинација неких од тих области), али са фокусом на примену савремених рачунарских технологија у тим областима.

Студирање на овом студијском програму подразумева да су студенти већ стекли одређена знања из неке од споменутих области друштвених наука. Такође се подразумева да студенти већ поседују елементарно познавање рачунарских технологија и да су упознати са основама коришћења рачунара и Интернета. Кроз студијски програм мастер академских студија Рачунарство у друштвеним наукама студенти се упознају са неким од напреднијих техника коришћења рачунара и Интернета, али увек кроз примену у одабраној области друштвених наука. Широка лепеза понуђених предмета омогућује сваком студенту да се усмери ка оној области друштвених наука која га највише интересује.

1. Кратак приказ програма

Овај једногодишњи програм носи 60 ЕСПБ, а осмишљен је тако да студенти прате курсеве у пролећном семестру (обично 5 курсева, који укупно носе 30 ЕСПБ), да би после радили на свом приступном раду (10 ЕСПБ) и мастер тези (20 ЕСПБ). Дипломирани студенти овог студијског програма добијају титулу Мастер рачунарства у друштвеним наукама.

Програм је осмишљен у складу са великим бројем сличних програма на европским и северноамеричким универзитетима. Зато је веома усмерен на квантитативне дисциплине и статистику, које су испреплетане са бројним темама из широких области друштвених наука и рачунарства. Два курса су обавезна:

- Савремене рачунарске технологије;
- Квантитативно моделовање у друштвеним наукама.

Они су предуслови за широк спектар изборних курсева:

- Анализа друштвених мрежа;
- Анализа и визуализација података;
- Демографија и нове информационе технологије;
- Методологија истраживања и статистика;
- Увод у когнитивну лингвистику;
- Дигитална хуманистика;
- Дигитализација и трансдисциплинарност у хуманистичким наукама;
- Програмирање за лингвисте;
- Правни и етички аспекти ИКТ;
- *Cyber*-криминал;
- Квантитативне методе у економији;
- Методе предвиђања и одлучивања;
- ИКТ и одрживи развој;
- Мерење информационог друштва;
- Анализа инвестирања у хартије од вредности;
- Технолошки подржано учење;
- Рачунарство и музика;
- Дигиталне библиотеке;
- Примена информационе теорије на обраду језика;
- Социјална психологија *Cyber* простора;
- Примене теорије узорковања у друштвеним наукама;
- Регулисање садржаја и слобода изражавања на Интернету;
- Развој људских ресурса и ИКТ;

- Управљање знањем и ИКТ;
- Управљање ризиком у актуарству;
- Примена информационих система у финансијској математици;
- Дигитализација и традуктологија.

Приликом одабира изборних предмета не постоје формална правила, али од студената се очекује да покажу сналажење у ширем пољу које је у вези са изборним предметом који су одабрали.

Програм је покренут фебруара 2015. године. Премда у време писања овог прегледа (почетком септембра 2015. године) ниједан студент још увек није завршио овај програм, прва генерација студената је одслушала све курсеве и већина је положила све испите. Због тога је било могуће да урадимо процену програма (крајем пролећног семестра) и да дођемо до релевантних статистичких података.

2. Кратак приказ програма

Једна од највећих предности овог програма коју је већина предавача који су учествовали у овом програму у пролеће 2015. године (око 50 предавача) нагласила јесте мотивација студената који су похађали предавања. Предавачима је пријало то што су студенти били вољни да уче, што су били вредни и веома активни за време предавања. Утисак већине предавача је био да су студенти могли да схвате нове теме веома брзо. То и није тако необично, с обзиром на то да су сви студенти прве генерације овог програма (36 студената) на основним студијама имали веома добре просечне оцене. С обзиром на чињеницу да многи студенти нису имали претходна знања из статистике и квантитативних дисциплина да је њихово познавање рачунарства у многим случајевима било ограничено на стандардне канцеларијске програме и Интернет претраживаче, сви одржани курсеви протекли су без потешкоћа.

Очигледан недостатак била је хетерогеност претходног знања студената, што је условило поделу студената у неким одељењима на две групе (потпуни почетници и они који имају бар нека искуства у вези са темом курса). То је, срећом, било надокнађено ентузијазмом студената и предавача, као и чињеницом да су сваки курс држала бар два предавача.

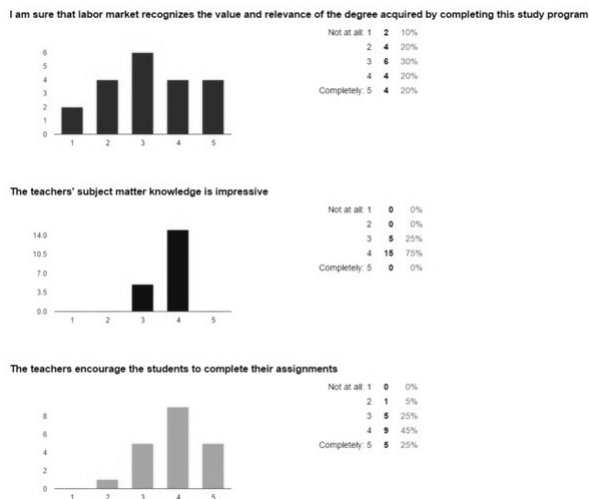
3. Из угла студената

У online упитнику за евалуацију који су студенти попуњавали на крају првог семестра било је питања о:

- квалитету понуђених курсева (о стеченом знању и томе колико га студенти сматрају корисним, о подударности са најављеним садржајем програма, о подударности са сличним мастер програмима на најбољим светским универзитетима (према виђењу студената), о коришћењу модерних алата и приступа учењу, материјала за учење и ресурса (литературе, софтвера ...) који су студентима доступни, online или у штампаном облику);
- сагледавању значаја дипломе која се добија завршетком овог студијског програма;
- виђењу квалитета предавања и педагошког рада које су показали предавачи;
- колико је програм прилагођен нивоу рачунарских предзнања студената који су завршили основне студије из друштвених наука;
- квалитету кабинета, опреме, Интернета, амбијента...;
- приказаном нивоу одзива и сарадње од стране предавача и административног особља;

Од укупно 36 студената попуњен упитник је предало 20 студената. На сва питања је требало одговорити користећи Ликертову скалу од пет нивоа (где је 5 највиша оцена). Било је дозвољено и писање слободних коментара.

Било је лепо сазнати из упитника да се студентима програм допада. Средње вредности добијених одговора на сва питања су увек биле више од 2,5, а често и више од 3. Слика 1 илуструје неке од резултата евалуације.



Слика 1. Неки од одговора студената

Студенти су веома ценили чињеницу да програм нуди могућност стицања многих практичних рачунарских знања и за њих је то најзначајнија карактеристика програма. Зато се чини да је овај програм одговарајућа надградња ширег теоретског и практичног знања у различитим областима друштвених наука које су студенти стекли током основних студија. Могућност да науче како да користе разне софтверске апликације у различитим пољима интересовања била је најважнији резултат. Студенти су, уопштено гледано, били веома задовољни показаним знањем предавача и углавном су били задовољни њиховим охрабривањем да заврше своје студентске задатке. Ова два фактора су умногоступно допринела успеху програма. Додатни позитиван подстицај била је чињеница да су студенти били веома активни, високо мотивисани и конкурентни.

Важно је нагласити и то да су студенти били веома задовољни различитошћу изборних курсева који су им понуђени. То им је омогућило да прилагоде избор курсева сопственим интересима и да искористе могућност да стекну нова знања и вештине у областима које су релевантне за њихову будућу каријеру. Основно познавање рачунарства данас се сматра неопходним на тржишту рада, док напредније вештине које нуди овај програм могу представљати потенцијалну предност у односу на друге дипломце из области друштвених наука.

Неким студентима је можда препреку представљао веома интензиван ритам програма, посебно током првог семестра. Да би се максимално повећала могућност стицања нових вештина и знања, изузетно је важно да студенти испуне све обавезе и да редовно похађају предавања, а она су се одржавала често. Зато само високо мотивисани студенти треба да размисле о уписивању на овај изузетно интересантан и користан постдипломски програм.

4. Шта смо научили?

Све у свему, резултати евалуације и искуства студената и предавача после првог семестра су позитивни, али има места за побољшања. На пример, из упитника за евалуацију које су студенти предали испоставило се да нису били задовољни усклађеношћу тема које су биле заступљене на предавањима са онима које су биле најављене у опису програма. То указује на чињеницу да би предавачи требало боље да образложе представљање најактуелнијих тема, методологија и алата, уместо строгог придржавања плана рада. Исто тако, није било мишљење свих студената да ће на тржишту рада њихова новостечена диплома бити препозната као вредна и релевантна. Премда су творци програма блиско сарађивали са Привредном комором и српским удружењима послодаваца

током осмишљавања курсева, узимајући у обзир потребе тржишта рада, свакако је потребно више труда да би студенти потпуно схватили ове предности.

Донекле виши ниво флексибилности очекује се и од предавача, у смислу давања задужења студентима – неки студенти су сматрали да им је за неке задатке био потребан доста висок ниво претходног техничког знања (што је мали број ових студената имао, с обзиром на то да су завршили основне студије из неке од области друштвених наука).

5. Наставиће се...

Очекује се да ће већина студената прве генерације студијског програма *Рачунарство у друштвеним наукама* да дипломира до фебруара 2016. године, када ће почети предавања за нову генерацију студената. Тада ће бити интересантно направити истраживање у вези са темама њихових завршних радова (многи студенти су тренутно у процесу одабира тих тема), како бисмо боље разумели интересовања и очекивања студената.

У међувремену, атмосфера у вези са овим програмом студија остаје позитивна. Потенцијални студенти ступају у контакт са члановима Програмског савета у вези с детаљима и охрабрује чињеница да интересенти долазе са разноврсних програма основних студија из области друштвених наука, као и са различитих универзитета. Ради илустрације, овде дајемо коментар потенцијалног студента (који има диплому из филозофије) послат председнику Програмског савета заједно са упитом о програму:

Сматрам да у данашње време заиста има смисла уписати се на овакав мастер програм с обзиром на чињеницу да на високо котираним универзитетима, какав је Стенфорд, студенти филозофије похађају курсеве филозофије заједно са курсевима рачунарства, лингвистике, психологије. То је једини начин да се филозофија заиста примени.

Приказ ЕУРОЛАН 2015 летње школе из Рачунарске лингвистике

Јелена Митровић
jmitrovic@gmail.com,
Универзитет у Београду
Филолошки факултет

РАД ПРИМЉЕН: 10. фебруар 2016.
РАД ПРИХВАЋЕН: 14. фебруар 2016.

ЕУРОЛАН летња школа, одржана од 13. до 25. јула 2015. године је била дванаеста у низу летњих школа које се сваке две године одржавају у Румунији. Теме ових летњих школа увек су веома актуелне, па је и ове године за тему одабрана област Рачунарске лингвистике и уопште управљања подацима и знањем – Лингвистички повезани отворени подаци (енг. Linguistic Linked Open Data – LLOD). Локација овогодишње школе био је прелепи град Сибиу, у срцу Трансилваније, окружен планинама и предивном природом.

Двадесет истакнутих предавача који важе за највеће стручњаке из области Рачунарске лингвистике, а свакако из области LLOD, који су умногоме и допринели њеном развоју, држали су интензивне курсеве током две недеље трајања ове летње школе. Преподнева су била намењена теоретском делу упознавања са новим концептима, док су поподнева углавном била испуњена практичним радом, туторијалима и вежбама.

Учесници су дошли из Европе, али и из Кине и Аустралије. Једна од предности овакве врсте стицања нових знања и унапређивања професионалних вештина јесте и могућност повезивања и дељења искустава са колегама из целог света – што је свакако изузетно важно! Ово је прва летња школа коју сам похађала и изузетно ми је пријало то што сам имала прилику да две недеље проведем са људима који деле моју страст према лингвистици и коришћењу рачунарских технологија за обраду природних језика.

У основи свих предавања и туторијала на овој летњој школи у некој мери се говорило о Семантичком вебу. Семантички веб је пројекат израде универзалног медијума за размену информација постављањем докумената са значењем које рачунар може да обради на вебу. Главни циљ концепта семантичког веба јесте семантичка интероперабилност веб извора, те постојање инфраструктуре за машинску интерпретацију и закључивање о садржајима на вебу.

Оквир за описивање ресурса (Resource Description Framework – RDF) је концепт који је настао у потрази за ефикаснијим решењима за проналажење информација и један је од стандарда Семантичког веба. Представља општи метод за концептуално описивање информација – семантичких веза између

електронских извора. Састоји се из уређених тројки (триплета): Субјекат – Предикат – Објекат где је Субјекат RDF URI референца извора који описујемо, Предикат је RDF URI референца, семантичка веза, а Објекат је RDF URI референца, сам meta datum.

Кључна технологија Семантичког веба је и SPARQL (изговара се као енглеска реч „sparkle“, а рекурзивни је акроним за SPARQL Protocol and RDF Query Language). Овај језик је развијен посебно за претраживање RDF база података и представља W3C стандард. Остали важни елементи семантичког веба су свакако XML (eXtensible Markup Language) који одређује структуру података (RDF/XML), онтологије, то јест модели представљања знања или скупови дефиниција неких концепата и релација које између њих постоје, OWL (Web Ontology Language) који се користи за објављивање и дељење онтологија. Све ове технологије кључне су за функционисање парадигме LLOD.

Повезани отворени подаци (енг. Linked Open Data – LOD), који су основ за LLOD, према основним принципима које је 2006. године поставио Tim Berners-Lee, јесу подаци који: 1) Користе URIs (јединствене идентификаторе извора) као имена ствари; 2) Користе HTTP URIs везе како би се та имена могла пронаћи 3) Пружају корисне информације помоћу RDF и SPARQL стандарда; 4) Садрже везе према другим URI ради откривања што више ствари; 5) Податке треба отворити за коришћење преко отворених лиценци.

Графички приказ LLOD облака настао је на иницијативу радне групе за отворену лингвистику (енг. Open Linguistics Working Group)¹ и ова група га уређује и одржава, а све то у склопу мреже отвореног знања – OKFN (Open Knowledge Foundation Network)².

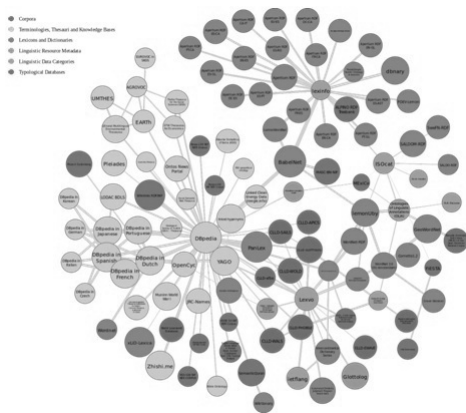
На слици 1 приказан је дијаграм на коме се види да су у LLOD облак укључени корпуси, базе знања, терминолошке базе, речници, лингвистичке категорије података, типолошке базе података.

DBpedia је веома важна у читавој LLOD и LOD парадигми. Она податке из Wikipedia страница трансформише у RDF. Садржи URI и друге метаподатке који се за сваку страницу формирају почевши од infobox делова Wikipedia страница. BabelNet је такође важан део LLOD облака. То је семантичка мрежа која обједињује WordNet, Open Multilingual WordNet (скуп свих отворених ворднетова), Wikipedia (највећа колаборативна енциклопедија), Wikidata (највећа колаборативна база знања), Wiktionary (највећи колаборативну речник), OmegaWiki (вишејезични речник средње величине).

Да би неки лексички ресурс био укључен у такозвани облак LLOD потребно је да буду испуњени следећи услови: 1) ресурс треба да буде доступан преко

¹ OWLG <http://linguistics.okfn.org/2011/05/20/the-open-linguistics-working-group/>

² OKFN <https://okfn.org/>



Слика 1. LLOD облак

разрешивих <http://> (или <https://>) URI веза; 2) подаци које садржи морају бити разрешиви у RDF податке у неком од најпознатијих RDF формата (RDFa, RDF/XML, Turtle, N-Triples); 3) Мора садржавати бар 1000 уређених триплета (енг. triples); 4) мора бити повезан преко RDF веза са неким ресурсом који је већ у LLOD дијаграму, или мора имати најмање 50 веза према неким другим ресурсима; 5) Може му се приступити RDF претраживањем веба (енг. RDF crawling), преко RDF dump-a, или преко SPARQL сервиса који прима захтеве и враћа резултате (енг. SPARQL endpoint).

На овој летњој школи сазнала сам и да се недавно родила идеја ширег прихватања LLOD технологије у WordNet заједници, то јест LLOD-a као основног механизма за креирање веза између ворднетова на различитим језицима, и то преко интерлингвалног индекса (енг. ILI). Прихватање отворених лиценци и дељених формата довело је до много доступнијих података из светских ворднетова. Тако смо имали прилике да се упознамо са Универзалним ворднетом који за циљ има решавање проблема полисемије и синонимије, а све кроз вишејезичност, то јест повезивање сличних или истих концепата на различитим језицима.

Закључак приче о Повезаним лингвистичким отвореним подацима јесте да су они корисно решење за многе примене јер: 1) пружају интеграцију информација – могуће је на ефикасан начин пронаћи и искомбиновати информације из различитих извора; 2) омогућавају вишејезичност и мултилингвалне примене многих алата; 3) динамично објављивање – подаци на вебу нису статични, могу се видети њихове различите верзије и исправити

грешке; 4) помоћу модела заснованих на графовима могуће је приказати било који облик језичког ресурса; 5) Проналажење информација је структурирано, на пример, можемо добити одговор на питање „Која су имена свих добитника Нобелове награде пореклом из Француске?“.

Поред непроцењивих знања о начину на који функционишу многе технологије семантичког веба и примера њихове практичне примене у конкретним пројектима, добила сам и много корисних предлога како бисмо српске лексичке ресурсе и алате могли да уврстимо у облак лингвистичких повезаних отворених података, и то од најбољих стручњака у тој области, од којих су неки поменуте технологије и осмислили, или значајно унапредили. Школа Еуролан је на мене свакако оставила веома позитиван утисак, а понајвише управо због инсистирања организатора да сви учесници проводе што више времена заједно, те су полазници курсева имали прилике да се боље упознају са предавачима и да од њих добију вредне савете. Следећа у низу ових летњих школа биће организована 2017. године и свакоме ко се бави Рачунарском лингвистиком бих топло препоручила да је похађа, а можда ћу то опет бити ја!

УПУТСТВО ЗА АУТОРЕ

Сви радови у часопису *Инфотека* објављују се на енглеском и на српском језику у истом издању. Аутори треба да доставе своје радове на једном од ова два језика; тек након обавештења о прихватању рада, очекује се да аутор пошаље превод (за српске ауторе; за све друге ауторе часопис ће обезбедити превод са енглеског језика на српски језик). Осим у штампаном издању, сви радови се објављују и онлајн у отвореном приступу.

КАТЕГОРИЗАЦИЈА РАДОВА

За документе прихваћене за објављивање и који подлежу рецензији примењује се следећа категоризација у часопису:

1. Научни чланци:
 - Оригинални научни рад (са претходно необјављиваним резултатима сопствених истраживања научном методом);
 - прегледни рад (који садржи оригиналан, детаљан и критички приказ истраживачког проблема или подручја у коме се допринос аутора показује и аутоцитатима);
 - претходно саопштење (оригинални научни рад прелиминарног карактера и мањег обима);
 - научна расправа и осврти на одређену тему заснована на научној аргументацији.
2. Стручни чланци који представљају искуства корисна за унапређење професионалне праксе
3. Информативни прилози могу бити:
 - уводници и коментари;
 - прикази књига, рачунарског програма, база података, стандарда и сл.;
 - научног догађаја, јубилеја.

Радови класификовани као научни морају имати бар две позитивне рецензије. Ставови Редакције не морају се подударати са ставовима у објављеним радовима. Рад се не може прештамповати нити објавити под сличним насловом нити у измењеном облику.

ЕЛЕМЕНТИ РУКОПИСА

За научне и стручне радове, потребно је доставити следеће податке:

1. Радови не би требало да буду дужи од 15 страна А4 формата, писаних врстом слова Times New Roman величине 12pt. У случају дужих радова, потребно је ступити у контакт са уредницима часописа.
2. Имена и презимена свих аутора требало би да буду написана оним редом којим ће се појавити у објављеном раду.
3. Пуно име једног или више аутора, без назнака титула и диплома, треба да буде достављено уз e-mail адресу, као и уз пун, званичан назив институције којој припадају (у сложеним организацијама се наводи укупна хијерархија назива, одозго надоле).
4. Наводи се и датум слања рада.
5. Аутори треба да предложи категорију свог рада, премда коначну одлуку о категоризацији доноси главни уредник.
6. Информативни сажетак, не дужи од 250 речи, који језгровито представља суштину рада, циљ истраживања и примењене методе и пружа главне закључке, треба да буде достављен уз рад. Сажетак треба написати на оба језика који се користе за објављивање часописа. У сажетку, аутори треба да користе термине који се као стандардни често користе за индексирање и претраживање чланака.
7. Аутори треба да доставе најмање 3 и не више од 10 кључних речи, одвојених зарезима, које означавају главне појмове представљене у раду. Списак кључних речи треба доставити на оба језика који се користе за објављивање часописа.
8. Ако рад произилази из мастер рада или докторске тезе, аутори треба да доставе назив тог рада или тезе, као и датум када су рад или теза предати и назив надлежне институције.
9. Ако рад представља резултат учешћа аутора у неком пројекту или програму, аутори би требало да укажу захвалност институцији која је била задужена за финансирање у посебном одељку под називом „Изјаве захвалности” на крају рада, пре одељка под називом „Литература“. У истом том одељку требало би навести имена особа које су помогле у изради рада.
10. Ако је неки рад представљен на конференцији, али није изашао у зборнику радова са те конференције, то такође треба нагласити у засебној белешци.
11. Аутори могу да користе фусноте, док употреба напомена на самом крају текста није дозвољена; ипак, треба избегавати употребу предугачких фуснота. Аутори могу да убаце додатке у своје радове.
12. Реферисани материјал треба да буде излистан у одељку под називом „Литература“ на крају рада. У списку реферисаних радова аутори треба да наведу све информације које су неопходне за проналажење радова на које се реферише. У овом одељку треба да се наведу све библиографске јединице реферисане у раду – такође, овде не треба да се појави ниједна библиографска јединица која у раду није реферисана.

ПРАВИЛА УРЕЂИВАЊА ПРИХВАЋЕНИХ РАДОВА

1. Препоручује се и подржава слање радове припремљених коришћењем пакета **ЛТ_ЕX**, коришћењем стила часописа (стил часописа и сви потрени пакети могу се преузети са сајта часописа). Аутори који нису упознати са коришћењем овог пакета треба да припреме свој рад у програму Word (формати .docx, .doc, .rtf или .txt). Аутори који користе програм Word на треба на посебан начин да га форматирају – пребацивање у **ЛТ_ЕX**е обавити редакција часописа.
2. Радови писани на српском језику треба да буду откуцани **ЋИРИЛИЧНИМ** писмом јер ће на том писму бити и штампани. Изузетак су једино делови текста за које је коришћење другог писма, на пример латиничног, погодније. Све врсте писама треба представити користећи Unicode UTF-8 кодни распоред.
3. Наслов рада не би требало да буде написан великим словима. Дужина наслова треба да буде у разумним границама – пожељно је да то буде мање од 150 карактера. За наслове дуже од 90 карактера, аутори треба да доставе скраћене верзије тих наслова које ће бити коришћене у заглављима.
4. Искошена слова (Italic) је могуће користити за истицање у тексту, док се подебљана слова (Bold) или искошена подебљана (Italic bold) могу користити ако је то неопходно. Употребу подвучених слова (Underlined) треба избегавати. Молимо да не истичете целе реченице, нити целе пасусе.
5. Рад може бити подељен у одељке и пододељке, али не треба користити више од два нивоа наслова одељака. Сви одељци ће бити на одговарајући начин нумерисани. Додаци треба да дођу на крај рада, а ако их има више и они ће бити означени. Ако се користе листе, не треба користити више од два нивоа угњеждености.
6. Сви пасуси треба да буду раздвојени једним празним редом (једним притиском на дугме Enter).
7. Код припреме табела и слика треба водити рачуна да се радови штампају у А5 формату, тако да треба избегавати сувише широке табеле. Све илустрације треба припремити у неком од формата који сажимањем не губе на квалитету, на пример .png, .tif или .jpg и у резолуцији од најмање 300 dpi.
8. Молимо ауторе да, ако је то могуће, додају везу на екран са кога је неки снимак екрана узет. Препоручујемо да аутори приликом чувања снимка екрана, или дела екрана, користе Zoom опцију претраживача или неког другог програма. За дијаграме који су произведени помоћу програма Excel, молимо да доставите оригинални .xls документ.
9. Све табеле, илустрације, дијаграме и фотографије треба припремити као засебне датотеке, црнобеле за штампање, а у боји за онлајн верзију. Наслови испод табела, илустрација, дијаграма или фотографија треба да остану у

тексту. Свака датотека треба да носи исто име као главни текст, уз опис врсте материјала коме је додат редни број у тексту. На пример, датотека у којој се налази четврта по реду слика у раду под називом „Пример“ треба да буде именована Пример_слика_4.

10. Молимо да додате све потребне датотеке у којима су објашњени неки посебни аспекти форматирања Вашег рада.
11. Уколико се у раду појављују URL адресе веб страница на чији се садржај текст реферише, њих би требало наводити у фуснотама, уз обавезно навођење датума када је страници приступљено.

БИБЛИОГРАФИЈА И ЦИТИРАЊЕ

1. Коришћену литературу навести на крају текста, у оквиру ненумерисаног одељка Литература. Овај одељак треба да садржи сва дела наведена у тексту, и ништа више од тога. Реферисану литературу не треба наводити у фуснотама у оквиру самог текста.
2. Сва дела треба сложити по азбучном реду имена аутора, уредника или издавача (ако аутор није наведен), а уколико се јавља више дела истог аутора њих треба поређати у хронолошком редоследу.
3. За израду референци користити Чикаго стил – (Chicago Style) (www.chicagomanualofstyle.org).
4. Наводити пуне, а не скраћене наслове часописа или акрониме.
5. Сви аутори, без обзира да ли предају рукопис припремљен коришћењем L^AT_EX-а или Word-а, припремиће све библиографске јединице из одељка литература и у BibTeX према шаблонима који су наведене уз примере на сајту Инфотеке (<http://infoteka.bg.ac.rs/index.php/sr/upu-s-v-z-u-r>).