

Импресум

ЗА ИЗДАВАЧА:

проф. др Александар Јерков

Универзитетска библиотека „Светозар Марковић“

Филолошки факултет, Универзитет у Београду

office@unilib.bg.ac.rs

ИЗДАВАЧ:

Филолошки факултет, Универзитет у Београду

Универзитетска библиотека „Светозар Марковић“

Заједница библиотека универзитета у Србији

ГЛАВНИ И ОДГОВОРНИ УРЕДНИК:

проф. др Цветана Крстев

Филолошки факултет, Универзитет у Београду

cvetana@matf.bg.ac.rs

ОПЕРАТИВНИ УРЕДНИК:

др Александра Тртовац

Универзитетска библиотека „Светозар Марковић“

aleksandra@unilib.rs

УРЕДНИК ЕЛЕКТРОНСКОГ ИЗДАЊА:

др Јелена Андоновски

Универзитетска библиотека „Светозар Марковић“

andonovski@unilib.rs

УРЕЂИВАЧКИ ОДБОР:

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

ISSN 1450-9687 (штампано издање) Београд, год. 21, бр. 1, септембар 2021.
ISSN 2217-9461 (е-издање)

ИЗРАДА ВЕБ ПОРТАЛА ЧАСОПИСА:

др Јелена Андоновски

Универзитетска библиотека „Светозар Марковић“

ЛЕКТОР ЗА ЕНГЛЕСКИ ЈЕЗИК:

Тања Ивановић

Министарство за европске интеграције

ДИЗАЈН И ПРИПРЕМА ЗА ШТАМПУ:

др Бранислава Шандрих Тодоровић

Филолошки факултет, Универзитет у Београду

ИнфоШека тим

РЕДАКТОР БИБЛИОГРАФИЈЕ И УДК КЛАСИФИКАЦИЈА:

Наташа Дакић

Универзитетска библиотека „Светозар Марковић“

РЕДАКТОР ДОО БРОЈЕВА:

др Милош Утвић

Филолошки факултет, Универзитет у Београду

РЕДАКЦИЈА ЧАСОПИСА:

Часопис ИнфоШека

Булевар краља Александра 71, 11000 Београд

+381 11 3370-211
infothecajdh@gmail.com

ШТАМПА:

Мамиго плус

Београд

Часопис излази два пута годишње

Садржај

Научни радови

- Лексичка база Фрејмнет: неколико примера
оквира из домена ризика**
А. Марковић, Р. Станковић, Н. Томић, О. Китановић 7
- О Корпусу студената англистике (КорСАНг)
и могућностима његове софтверске експлоатације**
М. Радоња, С. Шућур 37
- Вики-библиотекар, пројекат за обуку библиотекара
и студената за рад на Википедији**
Ђ. Стакић, А. Поповић, О. Кринуловић 59

Стручни радови

- Друштвени медији и њихова улога у ширењу
одређене представе о лепоти**
А. Сидики 79
- Инфотека (Q25460443) у Википодацима**
Р. Станковић, Л. Давидовић 93
- Мултимедијални пројекат Њих двоје**
К. Главоњић, М. Лукић, Ј. Јанковић, С. Јоксимовић 105
- EUROLAN 2021: Увод у повезане податке
у лингвистици – онлајн школа за обуку
полазника**
М. Дојчиновски, Х. Б. Хил, Х. Грасија, Р. Станковић 121

Прикази

- COBISS Meet 2020 - онлајн конференција Презентација
нових могућности за све учеснике мреже
COBISS.net**
А. Васиљевић, Ј. Зељић 129

Лексичка база Фрејмнет: неколико примера оквира из домена ризика

УДК 81'322.2

САЖЕТАК: У раду се даје кратак приказ теорије семантике оквира (енгл. *Frame Semantics*), на којој је заснована лексичка база Фрејмнет (енгл. *FrameNet*). Представљена је концепција ове мреже, као и могућности њене примене. Представљена је и лексичка анализа која се примењује у пројекту израде Фрејмнета и указано на разлике између анализе засноване на оквиру у односу на анализу засновану на речи. Затим је приказано неколико повезаних оквира које призивају речи из домена ризика. У раду је представљена и платформа NLTK (енгл. *Natural Language Toolkit*), помоћу које се могу користити разни језички ресурси, међу њима и Фрејмнет. Завршно поглавље пружа анализу именице *ризик* на корпусу рударства. Представљени су најчешћи колокати ове именице, скица њене употребе, конкорданце за поједине моделе, проналажење синонима и повезаних речи у виду тезауруса, графички приказ фреквенција појединих колокација, као и облака речи.

КЉУЧНЕ РЕЧИ: Српски језик, семантика оквира, Фрејмнет, сценарио ризика, корпус рударства, обрада природних језика.

РАД ПРИМЉЕН: 15. јули 2021.

РАД ПРИХВАЋЕН: 6. септембар 2021.

Александра Марковић

aleksan-

dra.markovic@isj.sanu.ac.rs

Институт за српски језик

САНУ

Београд, Србија

Ранка Станковић

ranka.stankovic@rgf.bg.ac.rs

Универзитет у Београду

Рударско-геолошки факултет

Београд, Србија

Наталија Томић

ntomic@hotmail.com

Универзитет у Београду

Београд, Србија

Оливера Китановић

olivera.kitanovic@rgf.bg.ac.rs

Универзитет у Београду

Рударско-геолошки факултет

Београд, Србија

1. Увод

Теорија семантике оквира (енгл. *Frame Semantics*) је когнитивносемантички приступ Чарлса Филмора који експлицитно

повезује значења неке речи са синтаксичким контекстима у којима се та реч јавља (Atkins, Fillmore, and Johnson 2003, 254). Анализом значења речи баве се углавном лексикографи, као и они који се баве семантиком. Међутим, уколико је циљ представљање начина на који се нека реч заиста користи, анализа података из корпуса показује се као прилично компликован задатак, с обзиром на број конкорданци које савремени корпуси нуде за поједине кључне речи. Теорија семантике оквира, према оцени више аутора (Atkins 1994; Gildea and Jurafsky 2002; Atkins, Fillmore, and Johnson 2003; Pradhan et al. 2005; Boas and Dux 2017; Jurafsky and Martin 2020), представља поуздан, научно утемељен начин да се анализира и опише начин употребе неке речи. У основи овог приступа лежи идеја да се свако искуство које памтимо јавља у неком смисленом контексту и да га можемо и упамтити баш захваљујући томе што имамо когнитивну схему или оквир за разумевање искустава. Филмор сматра да се речи уче у таквим смисленим контекстима, а ти су контексти неопходни и за процес разумевања, када у сећању призивамо искуства кроз која смо научили поједине речи. Оквир идентификује тип искуства, даје му структуру и целовитост, односно даје значење предметима, догађајима и односима у том искуству (Fillmore 1976, 26).¹

1.1 Концепција мреже Фрејмнет

Фрејмнет² је лексичка база података енглеског језика, заснована на анотацији примера употребе неке лексичке јединице (у даљем тексту ЛЈ; енгл. *lexical unit*) у аутентичним текстовима (дакле, у корпусу, а не у конструисаним примерима). Основна идеја своди се на то да се значења већине речи најбоље могу разумети на основу семантичког оквира, појмовне структуре налик на сценарио, који представља опис типа ситуације, догађаја, односа или ентитета, као и учеснике тих ситуација, догађаја и односа (Ruppenhofer et al. 2016, 7). На пример, ризиковање у типичном случају подразумева следећи скуп појмова: особу која је централна за сценарио РИЗИКА – *Протагонисту*, који свесно или несвесно доноси одлуку да ризикује или подлеже неком

1. Термин *оквир* Филмор користи као општи термин за скуп појмова који су у литератури о разумевању природних језика познати под називима *схема*, *сценарио*, *когнитивни модел*, *народна теорија* и неки други (Fillmore 1982, 111).

2. Пројекат се развија у Међународном институту за рачунарске науке у Берклију (International Computer Science Institute, Berkeley) од 1997. године.

ризик; могући *Лош исход* или штету; *Одлуку* која може довести до лошег исхода; *Циљ* који се жели постићи; *Околности* у којима се ризик јавља и *Вредност*, особу или предмет који су драгоцени Протагонисти и који су угрожени у датој ризичној ситуацији (Fillmore and S. Atkins 1994, 367).

1.2 Лексичка анализа заснована на оквиру

Лексичка анализа заснована на теорији семантике оквира подразумева анализу семантичког садржаја неке лексичке јединице, идентификацију њеног семантичког суседства, откривање граматичких конструкција у којима се јавља у корпусу, као и бележење свих конструкција у којима дата ЛЈ изражава пун семантички потенцијал. Сматра се да је неопходно описати све допуне и одредбе у тим конструкцијама. Посебна пажња поклања се речима које се не могу правилно употребити уколико се не знају конструкције у којима се јављају. Такве речи називају се речима које призивају оквире³ (енгл. *frame-evoking words*), а то су на првом месту глаголи, али и именице, придеви и прилози (Atkins, Fillmore, and Johnson 2003, 252).⁴

Основне јединице лексичке анализе у Фрејмнету су оквир и ЛЈ, под којом се подразумева лексема у једном од својих значења (Fillmore et al. 2003, 297), (Ruppenhofer et al. 2016, 7).⁵ За разлику од онога што је уобичајено у лексикографији, фокусирања на једну реч, односно лексему, и истраживања свих њених значења, тј. ЛЈ, у Фрејмнету се

3. Исти еквивалент енгл. глагола *to evoke* користи се и у преводу рада Fillmore 1982 у (Rasulić and Klikovac 2014, 79).

4. Теорија семантике оквира подстакла нас је да укажемо на потребу да се за четири основне класе речи које призивају оквире (именице, придеве, глаголе и прилоге) наводе релевантне конструкције у описним речницима српског језика (Марковић 2017, 34–41).

5. У српској лексиколошкој литератури, као и у синтаксичким радовима који се баве односом граматике и речника, користи се различита терминологија за оно што се у Фрејмнету назива лексичком јединицом (нпр. у универзитетском уџбенику лексикологије под лексичком јединицом подразумева се одредница или лема (Драгићевић 2007, 30)), док Поповић (2003, 202–203) указује на важност тога да се пажња у синтакси поклања конкретним лексемама, односно лексемама употребљеним у једном од више значења, које назива сублексемама. Ми смо се у овом раду определили за термин лексичка јединица, у складу с терминологијом мреже и приступа који представљамо.

у истом тренутку анализирају ЛЈ које припадају истом оквиру⁶ (Fillmore et al. 2003). Тако су, на пример, у односу на оквир *Бити угрожен* (енгл. *Being at risk*) дефинисане именице *ризик*, *опасност*, *безбедност*, *рањивост*, придеви *несигуран*, *безбедан*, *сигуран*, *поуздан*, *рањив* и др.

Процес описивања ЛЈ у Фрејмнету представљен је у (Fillmore et al. 2003).⁷ Он започиње неформалним описом оквира ком та јединица припада, прецизније описом типа ситуације или збивања представљених оквиром и стварања листе речи које се могу објаснити повезивањем с датим оквиром (299).⁸ Затим се бира циљна ЛЈ, централни члан оквира (у терминологији Фрејмнета *target*), ЛЈ у односу на коју се ради анотација; то је типично једна реч, али може бити и вишечлана реч или фразеологизам (Ruppenhofer et al. 2016, 21) и испитује њена употреба екстраховањем из корпуса реченица које је садрже.

Свој увид у значења изабране лексеме, добијен на основу посматрања примера из корпуса, лексикограф који ради на мрежи Фрејмнет пореди са значењима дате лексеме у неком од референтних речника.⁹ Када стекне јаснију представу о значењима, покушава прецизније да опише оквир у који спада изабрана ЛЈ. Затим пише дефиницију оквира – схематски опис типа ситуације која лежи у основи неке речи, заједно с називима улога учесника описане ситуације, које се називају елементима оквира (енгл. *frame elements*). У лексикографском смислу важан је начин на који елементи оквира добијају језички израз у реченицама у којима се јавља посматрана ЛЈ (Fillmore et al. 2003, 304–305).

6. Овде је у питању разлика између два приступа издвајању значења, први је заснован на речи (*word-based*), а други на оквиру (*frame-based*) (Atkins, Fillmore, and Johnson 2003, 254).

7. Иако се процес описује као низ одвојених корака, који се одвијају одређеним редом, аутори напомињу да се заправо у сваком тренутку може вратити на неки од претходних корака и ревидирати нека од одлука (Fillmore et al. 2003, 299).

8. Тај опис подразумева: 1) схематски опис типа ентитета или ситуације представљене оквиром; 2) бирање описних ознака за обележавање елемената оквира; 3) писање радне листе речи које припадају датом оквиру (припадност лексичких јединица истом оквиру указује на то да реченице које садрже те ЛЈ подлежу сличној семантичкој анализи) (297).

9. Пошто су анализи подвргли дефиниције глагола *ризиковати* у десет општих речника енглеског језика, Филмор и Еткинсова закључили су да чак ни речници сличних величина и намена не издвајају основна значења овог глагола, која припадају основном речничком фонду (Fillmore and S. Atkins 1994, 353).

1.3 Елементи оквира

Елементи оквира често се могу посматрати као примери општијих семантичких улога (као агенс, доживљавач, пацијенс), али се дефинишу на начин који је специфичан за поједине оквири (енгл. *frame-specific*). Постоји више разлога за то, а најважнији је тај што омогућава да се направе прецизне дефиниције елемената оквира за одређену групу речи, без потребе за утврђивањем начина на који ће се оне уклопити с малим, унапред задатим скупом општијих семантичких улога (305).

Најпре се утврђују централни елементи оквира (енгл. *core frame elements*).¹⁰ Централни елементи су они који представљају обавезне компоненте оквира и који неки оквир чине јединственим и различитим од других. Поред централних, постоје и периферни елементи, који се могу појавити у свим оквирима у којима неки агенс врши неку радњу (периферни елементи означавају појмове МЕСТА, ВРЕМЕНА, НАЧИНА, СРЕДСТВА).¹¹ Дешава се да неки од централних елемената не буде и језички изражен, иако је обавезан у појмовној структури оквира; такав случај назива се нултим појављивањем (енгл. *null instantiation*) и аотира се у бази (320). Пошто се одреде централни и периферни елементи, дефинише се оквир.¹²

Филмор и Еткинсова закључили су, након поменутог анализе глагола *ризиковати* у референтним речницима енглеског, да се допунама не

10. Постоје неке формалне особине које помажу у одређивању статуса елемента као централног (нпр. централни су они елементи који морају бити изражени, као и они који, уколико остану неизражени, добијају одређену интерпретацију (нпр. у реченици *John arrived* није изражено место на које је Џон стигао, као циљна локација, GOAL, али се то место разуме на основу контекста) (Ruppenhofer et al. 2016, 23–24).

11. Дистинкција централни/периферни елементи оквира одговара у грубим цртама подели на аргументе и одредбе у традиционалној граматичкој анализи (Fillmore et al. 2003, 310). Периферни елементи не могу бити у функцији субјекта или објекта циљног глагола и често су изражени прилозима или предлошко-падежним конструкцијама (319).

12. Ruppenhofer et al. (2016, 65) издвајају и додатне типове елемената оквира: они који се јављају у зависним клаузама имају периферни или екстра-тематски статус (енгл. *extra-thematic*, нпр. ВРЕМЕ, РАЗЛОГ). Осим тога, постоје и елементи којима се приписује ознака *Core-Unexpressed*; то су они елементи неког оквира који се понашају као централни неизражени, али који се можда неће наћи међу елементима оквира потомака датог оквира (24–25). У овом раду нећемо се бавити ни једним ни другим.

поклања довољно пажње (иако су веома битне, посебно у речницима за учење језика), а да постоје и неки други реченични чланови које речници занемарују, а који морају бити јасно издвојени и описани како би се посматрани глагол употребио правилно. На пример, радња коју изводи особа која нешто ризикује (и која може имати различите синтаксичке реализације): Ризиковала је свој живот *покушавајући да спасе дете које се давило*; затим, циљ који онај који ризикује има на уму кад се излаже ризику: Ризиковала је свој живот *како би спасла мој* (Fillmore and S. Atkins 1994, 362). Радња којом неко нешто ризикује спада у централне елементе оквира, док циљ ради ког се ризик предузима спада у периферне.

1.4 Односи између оквира – мрежа оквира

Пошто се дефинишу оквир и његови елементи, дати оквир се повезује са другим оквирима. На тај начин оквири, њихови елементи и с њима повезане ЛЈ добијају место у семантичком простору (Ruppenhofer et al. 2016, 79) и чине мрежу. Успостављање односа између оквира омогућава уочавање и бележење семантичких генерализација на основу типова учесника, догађаја и сл. Оквир може бити повезан са оквирима које разрађује или наслеђује, са онима који су његови подоквири, као и са онима које користи. Односи између оквира су усмерени или асиметрични: наиме, апстрактнији и независнији оквир се назива *Над_оквиром* (енгл. *Super_frame*) оквира који је зависнији и мање апстрактан и који се назива *Под_оквиром* (енгл. *Sub_frame*) (79).

Дефинисано је неколико врста односа, од којих су најважнији следећи (79–84):

- У односу *Наслеђивања* (енгл. *Inheritance relationship*) подоквир (подређени оквир) представља прецизнију разраду надоквира (надређеног оквира). Сви елементи оквира и подоквири надређеног оквира имају своје парњаке у подређеном оквиру, а он може имати додатне подоквири, елементе оквира и ограничења семантичког типа која се не јављају у Родитељу (Fillmore et al. 2003, 311). Нпр. оквир *Изложити се ризику* (*Run_risk*) наслеђује оквир *Вероватноћа* (*Likelyhood*).
- Однос *Коришћења* (енгл. *Using*) постоји када се неки оквир на веома уопштен начин позива на структуру неког апстрактнијег, схематизованог оквира. На пример, оквир *Клађење* (*Wagering*) користи оквир *Изложити се ризику* (*Run_risk*); *Брзина* (*Speed*)

користи оквир *Покрет* (*Motion*); оквир *Брбљивост* (*Volubility*) користи оквир *Комуникација* (*Communication*) (Ruppenhofer et al. 2016, 83).

- *Угао_гледања* (енгл. *Perspective_on*) је релација слична општијој релацији *Коришћења*, али повезане оквири ограничава у већој мери (82). Примена ове релације указује на постојање бар два могућа угла гледања на један неутралан оквир. На пример, оквир *Сценарио_ризика* (*Risk_scenario*) је неутралан, док су оквири *Ризична_ситуација* (*Risky_situation*), *Бити_угрожен* (*Being_at_risk*) и *Изложити се_ризик* (*Run_risk*) перспективизовани; оквири *Запошљавање* (*Hiring*) и *Добити_посао* (*Get_a_job*) перспективизују оквир *Почетак_радног_односа* (*Employment_start*) са становишта послодавца, односно запосленог.

Пошто се у базу унесу дефиниције оквира и његових елемената, оквиру на ком се ради додаје се ЛЈ (на пример, оквиру *Бити_угрожен* додаје се ЛЈ *ризик*). Затим се за дату ЛЈ уносе информације о врсти речи, значењу, формалном саставу (једна реч или вишечлани израз). Након тога прецизирају се претраге које ће омогућити да се из корпуса¹³ екстрахују оне реченице (поткорпус) које садрже лексему коју испитујемо (у овом примеру именицу *ризик*) и чија је граматичка форма таква да указује на оно значење дате именице (ЛЈ) које је везано за оквир *Бити_угрожен* (*Being_at_risk*). Циљ је да се из поткорпуса избаце све оне реченице у којима задата кључна реч не представља ЛЈ која се везује за оквир који се креира. Пошто се спецификују претраге важне за дату ЛЈ, низ аутоматских процеса генерише поткорпус спреман за анотацију. Након што се одбаце предуге и друге неадекватне реченице, бира се од три до пет примера за сваки модел – циљ је илустровати разноликост модела, а не постићи статистичку репрезентативност.

Кад се заврши с анотацијом, примењују се алати за испитивање анотираних реченица и валентних модела који су у њима реализовани. Постоје два типа извештаја у виду динамичких веб-страница (*LexUnit Report*, *Lexical Entry Report*; извештаји се аутоматски генеришу пошто се заврши анотација и доступни су и на јавној веб-страни Фрејмнета). Први показује све анотиране реченице за једну ЛЈ и у њему се наводе сви елементи оквира пронађени у актуелном оквиру (табела елемената

13. Fillmore et al. (2003, 304) користе Британски национални корпус (енгл. British National Corpus).

оквира); сваки елемент се боји одређеном бојом, а истом том бојом елементи су обележени и у анотираним реченицама. Други представља резиме синтаксичких реализација елемената оквира и валентних модела ЛЈ у две табеле (Fillmore et al. 2003, 326–328).

Пошто се у Фрејмнету бележе и елементи оквира (за оквир специфичне семантичке улоге), као и њихове језичке реализације, за овај опис важни су и појмови *валентне групе*, *валентног модела* и *описа валенце*.¹⁴ Елемент оквира заједно са својом граматичком реализацијом (тип јединице и њена функција у реченици) чини валентну групу; скуп валентних група реализованих у једној реченици чини валентни модел; скуп свих валентних модела које реализује једна лексичка јединица чини опис валенце (Atkins, Fillmore, and Johnson 2003, 255–257).

1.5 Примене мреже Фрејмнет

Фрејмнет је доступан на адреси <https://framenet.icsi.berkeley.edu/>. Може се претраживати и прелиставати онлајн, али и преузети и користити. Као што је на сајту истакнуто, може се користити у различите сврхе: као речник за учење језика (с обзиром на то да садржи више од 13.000 ЛЈ); као речник валенце; као скуп података за обуку програма који обележавају семантичке улоге,¹⁵ што га чини значајним дигиталним језичким ресурсом за истраживања у области обраде природних језика (садржи више од 200.000 ручно обележених реченица повезаних са више од 1.200 семантичких оквира).

Фрејмнет је настао као лексичка база података за енглески, а касније су развијане базе за друге језике (француски, кинески, португалски, немачки, шпански, јапански и др.), кроз различите и независне пројекте, али уз исти формализам и концепцију развоја. Покренут је пројекат поравнавања датих база података за разне језике.

14. Својство глагола да за себе „везује“ зависне елементе (аргументе) назива се валенца. У зависности од броја аргумената које везују за себе, глаголи могу бити једновалентни (кад се могу употребити само са субјектом), двовалентни (са субјектом и објектом) итд.

15. У подтачки 1.6 представимо нека од истраживања у којима се на разне начине користе Фрејмнет и програми за бележење семантичких улога у хрватском, словеначком и српском.

1.6 Претходна истраживања

У овом одељку осврнућемо се на истраживања која су се бавила бележењем семантичких улога (анотација семантичким улогама) у српском и њему сродним језицима и испитивањем значења именице *ризик* и глагола *ризикувати*.

У раду Gantar et al. (2018) представљен је модел за обележавање семантичких улога у словеначком и хрватском, развијан у оквиру међународног билатералног пројекта *Semantic Role Labeling in Slovene and Croatian*. Циљ је био развијање ручно анотираних корпуса који би се користио као база података за обуку за системе надгледаног машинског учења. Описан је и експеримент аутоматског обележавања семантичких улога заснован на надгледаном машинском учењу. За сваки корпус представљени су најфреквентнији глаголи, семантичке улоге и типични семантичко-синтаксички обрасци за најфреквентније глаголе. У оба корпуса најфреквентнији је глагол *бити* и семантичка улога пацијенса, а затим агенса (95–96). У раду су семантичке улоге бележене на стабилним семантичко-синтаксичким моделима (96–97), али остаје питање оправданости таквог поступка будући да се семантичке улоге и оквири везују за ЛЈ, дакле за (глаголске) лексеме у једном од значења.

У раду Brač and Anić (2019) описан је пројекат чији је циљ развој методологије за бележење семантичких улога у језику специјализованих домена (конкретно авијације) који би се могао користити за све посебне области. Ауторке разматрају да ли је боље користити општије семантичке улоге или улоге специфичне за поједине глаголе и оквири, што је карактеристика Фрејмнета. Њихов закључак је да превелики број специфичних семантичких улога успорава анотацију а не доприноси значајно бољитку термилошких ресурса, док се скуп општијих семантичких улога мора незнатно допунити (545).

У раду Wasserscheidt and Hršić (2020) спроведено је интересно истраживање о томе да ли лексеме које се користе и као термини (конкретно правни термини) и као речи општег лексичког фонда имају различита значења (или призивају различите оквири) у српском и хрватском језику (као језичким варијантама). Идеја је потекла од контрадикције у литератури посвећеној семантици оквира; наиме, у изворној Филморовој варијанти теорије указује се на разлике у оквирима које постоје између појединих говорника, друштвених група и култура; на те се разлике у радовима других аутора заборавља, па се оквири посматрају као универзалне структуре, независне од језика (88–89). Аутори су испитивали да ли *одредба* као правни термин

и реч општег лексичког фонда призива различите оквири у српском и хрватском, а испитивање је спроведено на корпусима ових језика. Коришћена је дистрибуциона анализа, чија је основна идеја да се значење речи може разумети на основу њеног контекста, али је тој анализи подвргнут и сам контекст. У анализи контекста примењена је и теорија семантике оквира (Wasserscheidt and Hrstić 2020, 90). На основу ове двоструке дистрибуционе анализе аутори су закључили да нема значајне разлике у значењу *одредбе* у посматраним корпусима, као и то да се метод двоструког кластеровања може користити у сложеној семантичкој анализи, која се пак може уклопити у појмовне структуре Фрејмнета¹⁶ (108).

Иако није од непосредног значаја за наш рад (јер представљена анализа није заснована на теорији семантике оквира), поменућемо истраживање у ком се најпре примећује да је ризик постао важна тема истраживања у друштвеним наукама, али да се, када се говори о ризику, о значењу саме речи *ризик* не зна довољно (Hamilton, Adolphs, and Nerlich 2007, 164). Аутори истраживања, вођени поменутиим запажањем, подвргавају анализи значење именице *ризик* и глагола *ризиковати* у корпусу разговорног језика (*the Cambridge and Nottingham Corpus of Discourse in English*, скраћено CANCODE). Испитујући семантичку склоност поменутих лексема, као и њихову семантичку прозодију, аутори истраживања закључују да на анализиране лексеме утиче контекст у ком се јављају (нпр. постоје разлике између њихових колоката и семантичке прозодије у интимној комуникацији, између чланова породице или партнера, у односу на комуникацију између професора и студената).

2. Неколико примера оквира из домена ризика

Као што је поменуто и на крају т. 1.3 овог рада, Филмор и Еткинсова разматрали су ограничења која су лексичкој анализи наметнута у традиционалном приступу лексикографији и у папирним издањима речника (Fillmore and B. T. Atkins 1992, 100–101), (Fillmore and S. Atkins 1994, 350–363). Пошто су упоредили анализе глагола *ризиковати* и именице *ризик* у референтним једнојезичним речницима с налазима о овим лексемама до којих су дошли анализом корпуса, схватили су да

16. Анализа је указала на то да се *одредба* јавља у чак 12 различитих оквира (Wasserscheidt and Hrstić 2020, 108).

ниједан од прегледаних речника не пружа задовољавајућу анализу и не бележи многе податке које су уочили прегледом примера из корпуса. Закључак је био да папирни речник, у ком је приказ линеаран, не може да представи тако сложен опис какав би био потребан да би се представили сви подаци важни за употребу неке речи. Због тога су осмислили онлајн речник заснован на оквирима, уместо папирног речника заснованог на лексемама, у ком је могуће представити резултате овакве анализе.

Овако замишљен онлајн речник омогућава да се прикажу поједини елементи оквира и њихове различите синтаксичке реализације, дакле све оно што спада у опис валенце (коју смо поменули у т. 1.4), као и односи између различитих оквира.

Алат за визуелизацију односа између оквира и елемената оквира (*FrameGrapher*)¹⁷ пружа могућност да се изабере почетни оквир и систематски истраже везе између оквира. Сlike 1–4 у наставку текста генерисане су управо коришћењем алата *FrameGrapher*.

2.1 Оквир *Ризична_ситуација*

У наставку је представљен оквир *Ризична_ситуација*.¹⁸ Након дефиниције дају се примери, а потом централни и периферни елементи оквира. Као што је поменуто, сваком елементу оквира додељује се посебна боја, тако да се и у тексту дефиниције елементи, ако се појаве, боје одговарајућим бојама. Лексичке јединице које призивају оквир *Ризична_ситуација* су: *опасност.п*, *опасан.а*, *ризик.п*, *рискантно.adv*, *ризичан.а*, *безбедан.а*, *безбедно.adv*, *небезбедан.а*, *претња.п*, *шкодљив/несигуран.а* (*п* је скраћеница за именице, *а* за придеве, *adv* за прилоге, *в* за глаголе). ЛЈ које призивају оквир маркирају се црном позадином у анотираном тексту. За сваки елемент оквира наводи се дефиниција и један обележен пример употребе.

17. *FrameGrapher*

18. За потребе овог рада превели смо опис неколико оквира и њихових елемената, уз напомену да је реч о опису везаном за енглески, са примерима из корпуса енглеског језика, прилагођеним српском језику ради илустрације принципа. Надамо се да ћемо опис оквира заснован на подацима из корпуса српског језика представити ускоро.

Ризична ситуација

Дефиниција:

Одређена **Ситуација** може (али не мора) да доведе до штетног догађаја који би задесио неку **Вредност**. Та **Ситуација** може бити неко стање, активност или неки кључни ентитет који мора бити схваћен као део неке шире **Ситуације**, која укључује и тај ентитет и **Вредност**. Иако је за разумевање овог оквира кључна идеја о штетном догађају, он не мора бити изражен као аргумент лексичких јединица у овом оквиру.

Да ли су **климатске промене** **ОПАСНЕ** по човечанство?

Купци се могу жалити на **ШКОДЛИВЕ** **производе**.

Највећа **ОПАСНОСТ** прети **нашој** **инфраструктури**.

Елементи оквира

Централни:

Вредност (Asset)
[ass]

Нешто што се сматра вредним и пожељним, а постоји могућност да ће му штета бити нанета или да ће бити изгубљено.

Мед није **БЕЗБЕДАН** за **бебе**.

Опасан ентитет
(Dangerous entity)
[dan ent]

Конкретан или апстрактан ентитет који може нанети штету **Вредности** или довести до њеног губитка.

Том човеку је **ФИЗИК** друго име!

Искључује: Ситуацију

Ситуација
(Situation) [sit]

Ситуација може довести до неког штетног догађаја. Већина људи се слаже да није **БЕЗБЕДНО** возити брже од 120 km/h.

Периферни:

Околности
(Circumstances) [c]

Околности под којима је **Вредност** угрожена.

Степен
(Degree) [deg]

Одредба која изражава одступање тренутног нивоа безбедности од очекиване вредности, узимајући у обзир **Ситуацију** и стање на које указује циљна ЛЈ.

Терористи су наша **највећа** **ПРЕТЊА**.

Домен
(Domain) [dom]

Домен у коме је **Ситуација** безбедна.

Све наше сајтове је **БЕЗБЕДНО** користити у **образовању**.

Учесталост
(Frequency) [f]

Колико често **Вредност** долази у **Ризичну ситуацију**.

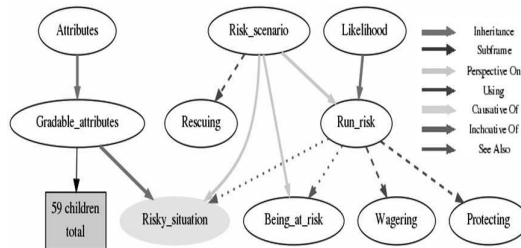
Место
(Place) [pla]

Одређена локација на којој је **Ситуација** безбедна.

Често се може закључити да карактеристике неке локације чине одређене **Ситуације** безбедним или небезбедним.

Време
(Time) [tim]

Временски период током ког одређена **Ситуација** има прецизирани ниво сигурности.



Слика 1. Графички приказ односа између оквира *Ризична_ ситуација* (енгл. *Risky_ situation*) и с њим повезаних оквира

2.2 Оквир *Бити_угрожен*

Оквир *Бити_угрожен* (енгл. *Being_at_risk*) призивају следеће лексичке јединице: *опасност.н*, *несигуран.а*, *ризик.н*, *безбедан.а*, *сигуран.а*, *безбедност.н*, *поуздан.а*, *рањивост.н*, *рањив.а*. Боје одговарајућих елемената оквира исте су као у претходном примеру; овај оквир садржи и додатни елемент – *Штетни_догађај* (*Harmful_event*).

Бити_угрожен

Дефиниција:

Вредност је у неком стану у ком је изложена или подложна дејству **Штетног догађаја**, који може бити метонимијски призван дејством **Опасног ентитета**. Речи које означавају релативну сигурност (одсуство ризика) такође су део овог оквира.

Нема делета које је **ЗАШТИЋЕНО** од искупења да уради оно што раде и његови вршњаци. **Наша држава** ужасно греши покушавајући да се **ЗАШТИТИ** од **зузи** – она мора да штити суде.

Уколико радим као багериста, **ви** си под **РИЗИКОМ** од губитка слуха због изложености **буци** доком радите.

Ви нисте **СИГУРНИ** од крађе података уколико немате заштиту од прислушкивања.

Елементи оквира:

Централни:

Вредност (Asset) [ass] Нешто што се сматра пожељним или драгоценим и што може бити изгубљено или оштећено. Закључани катанац гарантује да су **Информације СИГУРНЕ**.

Опасан ентитет (Dangerous entity)

Конкретан или апстрактан ентитет који може да узрокује губитак или оштећење **Вредности** због њеног учешћа у **Штетном догађају**. Старомо се да ваш **БМ** буде **БЕЗБЕДАН/ЗАШТИЋЕН** од **провалника**.

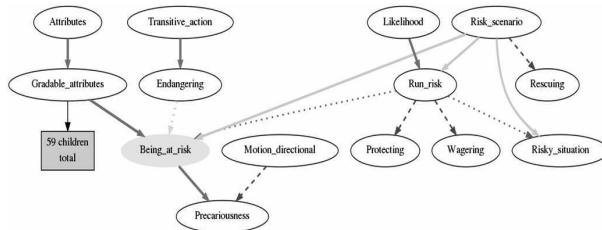
Штетан догађај (Harmful event) [har]

Догађај који се може одиграти или стање које се може одржати и које може довести до губитка или оштећења **Вредности**.

Искачуције:

Опасан ентитет (Dangerous entity)

Наш систем обезбеђује да информације које се чувају на хардверу буду **ЗАШТИЋЕНЕ** од напада хакера, као и од покушаја физичке крађе.



Слика 2. Семантички оквир *Бити_угрожен* (енгл. *Being_at_risk*)

2.3 Оквир *Изложити се_ризик*

Оквир *Изложити се_ризик* (енгл. *Run_risk*) призивају следеће лексичке јединице: *угрожен.а*, *опасност.н*, *ризик.н*, *ризиковати.в*, *угрозити.в*. На слици 3 дати су дефиниција, примери и елементи овог оквира.

Изложити се ризику

Дефиниција:

Протагониста је у потенцијално опасној ситуацији која може да се оконча Лошим исходом по њега или њу. Опасност од губитка Вредности може да представља Лош исход. Не постоје назнаке да се Протагониста намерно излаже ризичној ситуацији. Могуће је да Протагониста покушава да оствари неки Циљ, чиме доспева у опасну ситуацију. Завиљност ризика којем се излаже такође може бити изражена. Постојао је РИЗИК да се све запали.

Елементи оквира:

Централни:

Радња (Action) [Act] Радња која изазива ризик. Имплементација овог програма излаже нас РИЗИКУ да увредимо своје најверније бираче. Нешто пожељно што Протагониста поседује или је с њим непосредно повезано и што може бити изгубљено или оштећено. То је био велики РИЗИК по његову репутацију.

Вредност (Asset)

[Asset]

Искључује:

Лош исход

(Bad outcome)

Лош исход (Bad outcome)

Bad]

Протагониста

(Protagonist)

Периферни:

Бенефицијар

(Beneficiary) [ben]

Циљ

(Purpose) [Pur]

Завиљност

(Severity) [Sev]

Ситуација коју би Протагониста хтео да избегне.

РИЗИКОВАО је да изгуби своје здравље.

Особа којој прети неки Лош исход.

Протагониста има за циљ да његови поступци буду на корист Бенефицијару.

Све би РИЗИКОВАЛИ за своје најближе пријатеље.

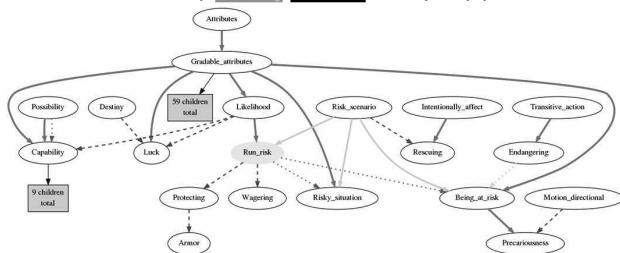
Жељена радња или циљ за који Протагониста верује да ће га остварити.

Сви су они РИЗИКОВАЛИ да буду ухапшени само да би се домогли Америке.

Вероватноћа да ће се нешто лоше десити

Протагонисти.

Када се више не разликује добро и зло, човек је у Завиљној ОПАСНОСТИ да изгуби душу.



Слика 3. Семантички оквир *Изложити се ризику* (енгл. *Run risk*)

Сценарио ризика

Дефиниција:

Вредност је у ситуацији за коју је вероватно да води до неког штетног догађаја, који ће лоше утицати на Вредност.

Елементи оквира:

Централни:

Вредност (Asset) [Ass] Нешто што се сматра пожељним или вредним и што може бити изгубљено или оштећено.

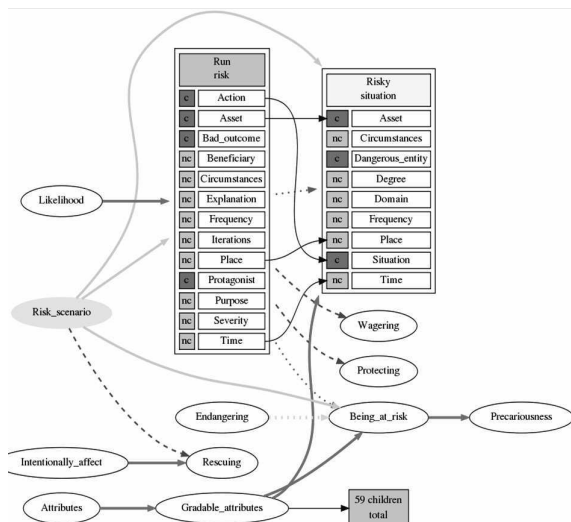
Штетни догађај (Harmful event) [Har] Догађај који може да се одигра или стање које може да потраје и које може довести до губитка или оштећења Вредности.

Ситуација (Situation) [Sit] Ситуација у којој Вредност није безбедна или заштићена.

Периферни: Степен (Degree) [Deg] Одредба која изражава одступање од актуелног нивоа безбедности за Вредност, ситуација и стање означени самом циљном јл.

Место (Place) [Pla] Место за које важи одређени ниво безбедности.

Време (Time) [Tim] Време током ког важи одређени ниво безбедности.



Слика 4. Семантички оквир *Сценарио ризика* (енгл. *Risk_scenario*)

2.4 Сценарио_ризика

Слика 4 представља оквир *Сценарио_ризика* (енгл. *Risk_scenario*), са детаљним приказом оквира *Изложити се_ризикy* (*Run_risk*) и *Ризична_ситуација* (*Risky_situation*), за које су обележени централни (енгл. *core*, скр. *c*) и периферни (енгл. *non-core*, скр. *nc*) елементи. На десној страни слике дата је легенда с врстама релација које се успостављају између оквира, на пример: наслеђивање, угао гледања, коришћење (приказане су и неке од релација које нисмо помињали: подоквир, узрочник и др).

3. NLTK омотач за Фрејмнет

NLTK (енгл. *Natural Language Toolkit*) је платформа за развој програма *Python* за обраду текста на природним језицима, која је једноставна за коришћење, а садржи бројне корпусе и лексичке ресурсе чији се број континуално увећава. NLTK омогућава различите врсте обраде текста, међу којима су: класификација, токенизација (енгл. *tokenization*), стемовање (енгл. *stemming*), тагирање (енгл. *tagging*), парсирање (енгл. *parsing*) и семантичко закључивање (енгл. *semantic reasoning*). У оквиру NLTK система имплементирани су и омотачи (енгл. *wrappers*) за друге програмске библиотеке за обраду природних језика, као и за важне лексичке ресурсе. Један од расположивих ресурса у оквиру NLTK је и Фрејмнет, који прати програмска библиотека за претраживање овог ресурса и екстракцију информација из његових оквира.

Као што је поменуто у Уводу (одељак 1.1 овог рада), оквир је појмовна структура¹⁹ која описује одређену врсту ситуације, ентитета или односа заједно са учесницима тих ситуација и односа. Физичка организација Фрејмнета у оквиру NLTK дистрибуције представља колекције XML (*Extensible Markup Language*) датотека разврстаних по каталозима: *frame*, *fulltext*, *lu*, *miscXML*, којима се приступа помоћу функција библиотеке или се могу директно претраживати и приказивати коришћењем XML докумената са XSL (енгл. *eXtensible Stylesheet Language*) трансформацијама: *frameIndex*, *luIndex*, *fulltextIndex*. У овом одељку ћемо приказати примере коришћења функција омотача Фрејмнета.

19. FrameNet и NLTK

Да би се добила листа свих оквира у Фрејмнету, може се користити функција `frames()`. Следећи код приказује иницијализацију рада са корпусом Фрејмнет и показује да верзија која је публикована уз NLTK има 1221 оквир.

```
from nltk.corpus import framenet as fn
len(fn.frames())
```

Да бисмо пронашли оквире који у називу садрже реч 'risk', користи се опција:

```
fn.frames(r'risk')
```

која као резултат даје:

```
[<frame ID=1560 name=Being_at_risk>,
 <frame ID=378 name=Run_risk>].
```

Имајући у виду осетљивост на мала и велика слова, за упит:

```
fn.frames(r'Risk')
```

добива се другачији резултат:

```
[<frame ID=1763 name=Risk_scenario>,
 <frame ID=1762 name=Risky_situation>]
```

Ако се зада регуларни израз '(?i)risk' функцији `frame()`, добија се листа од сва четири поменуто оквира (т. 2.1–2.4), чија се имена подударaju с прослеђеним обрасцем, јер '(?i)' указује на то да се не прави разлика између малих и великих слова.

Детаљи одређеног оквира могу се добити функцијом `frame()`, којој се проследи као параметар број оквира, рецимо `f=fn.frame(1762)`, што ће вратити податке за оквир Ризична ситуација (*Risky_situation*).²⁰

Уз то, може се прићи деловима оквира, рецимо називу оквира (`f.name`), његовој дефиницији (`f.definition`), елементима оквира (`f.FE`), лексичким јединицама (`f.lexUnit`), релацијама између оквира (`f.frameRelations`), што је илустровано следећим кодом:

```
f = fn.frame('Risky_situation')
print(sorted([e for e in f.FE]))
print([r for r in f.frameRelations])
```

20. Подаци за оквир *Ризична ситуација*

који даје следећи резултат:

```
[ 'Asset', 'Circumstances', 'Dangerous_entity', 'Degree', 'Domain',
  'Frequency', 'Place', 'Situation', 'Time' ]
[ <Parent=Gradable_attributes - Inheritance →
  Child=Risky_situation>,
  <MainEntry=Run_risk - See_also →
  ReferringEntry=Risky_situation>,
  <Source=Run_risk - ReFraming_Mapping →
  Target=Risky_situation>,
  <Neutral=Risk_scenario - Perspective_on →
  Perspectivized=Risky_situation> ]
```

4. Анализа речи *ризик* на корпусу рударства

Развој једнојезичног корпуса из домена рударства почео је као део пројекта везаног за управљање рударском пројектном документацијом коришћењем језичких технологија (Тomašević et al. 2018, 996). Тада је једнојезични корпус обухватао текстове из рударског домена и сродних истраживања, са укупно 172 документа (на српском језику) и 2,7 милиона речи у првом издању (997). Током даљих истраживања (Kitanović 2021) проширен је са додатна 63 документа. Тренутна верзија садржи 4,1 милион речи. Класификован према изворима, корпус обухвата пројектну документацију (26%), законодавство (11%), докторске дисертације (31%), уџбенике и осталу рударску литературу (32%) (Kitanović et al. 2021, 8).

Резултат претраге у CQL-y²¹ (енгл. *Corpus Query Language*) анализира се на различите начине: листа фреквенција, колокације, конкорданце са ужим и ширим контекстом. Слика 5 приказује конкорданце, добијене из веб-апликације Лексимирика²² за управљање електронским речницима (Stanković et al. 2018), за образац придев-именица за именицу *ризик*, док се на слици 6 може видети хистограм са фреквенцијама препознатих облика за исти образац, екстрахованих из рударског корпуса, доступног на корпусној платформи отвореног кода *NoSketch Engine* (Kilgariff et al. 2004).²³ Инстанцу на локалном серверу одржавају чланови Друштва за језичке ресурсе и

21. Corpus Querying

22. Лексимирика

23. NoSketch на JePTexy, NoSketch Engine

Concordances – Profil 1 – Microsoft Edge		
https://leximika.jerteh.rs/CQP/Noske		
A(N)		
na najmanju moguću meru, odnosno otklanjanje	profesionalnih rizika	. Strategija teži da se u ovom periodu broj
na najmanju moguću meru, odnosno otklanjanje	profesionalnih rizika	. Strategija teži da se u ovom periodu broj
inspektora rada sa novim tehnologijama i	novim rizicima	. savremenim pristupima i praksama u oblasti
i zdravља na radu uzimajući u obzir	posebne rizike	koji se pojavljuju u određenim delatnostima . o
uzajamna povezanost, što samo još povećava	potencijalne rizike	za ukupnu realizaciju procesa rekultivacije .
velikih količina otpada po konkurentnoj ceni a	niskom riziku	po životnu sredinu ; 2. ekonomski isplativ
. potencijalno moguće ozbiljnije povrede,	mali rizik	fatalnog kraja . gubici radnog vremena Nizak
i normalna komunikacija Neophodna prva pomoć,	mali rizik	od ozbiljnih povreda Zanemarujući Nemerljivi
6. , jer osim snabdevanja gasom , kod nje postoji i	izvesni rizik	od povraćaja investicije , što bi moglo
odgovora često veoma zahtevan , složen , i sa	prisutnim rizicima	. Konačno rešenje , kako smo već istakli u
i sl. • Prisustvo konfliktnih situacija ,	povišenih rizika	i nepovoljnih događaja , npr. interakcija
sistema zaštite na radu : 1) radnim mestima sa	povećanim rizikom	; 2) zaposlenima raspoređenim na radna mesta
: 2) zaposlenima raspoređenim na radna mesta sa	povećanim rizikom	i lekarskim pregledima zaposlenih
može da ima previd pojedinih opasnosti . Psiho -	socijalni rizici	se obično prevede , kao i rizici u vezi sa
, a takođe je ostvaren napredak i u proceni	profesionalnih rizika	i sistematizaciji profesionalnih bolesti .
ili smanjenja rizika . Radno mesto sa	povećanim rizikom	jeste radno mesto utvrđeno aktom o proceni
je da se nekontrolisane opasnosti prevedu u	kontrolisani rizik	i da se na taj način bolje zaštite zaposleni i
identifikovanju i kontrole zdravstvenih i	sigurnosnih rizika	organizacije i eliminisanju ili smanjivanju
organizacije i eliminisanju ili smanjivanju	potencijalnog rizika	od nezgoda na prihvatljivi nivo , poštujući pri
najvišeg mogućeg nivoa bezbednosti i	minimalnog rizika	moraju se dokumentovati uključujući i zapise

Слика 5. Конкорданце обрасца придев–именица за именицу *ризик*

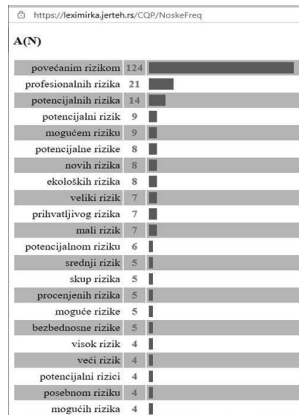
технологије JePTex.²⁴ За тагирање корпуса обучен је *Treegger* модел за српски (Krstev and Vitas 2005; Utvić 2011), (Stanković et al. 2020, 3957), коришћењем ручно анотираних корпуса и Српских морфолошких речника (Krstev 2008).

Рударски корпус је публикован и на платформи *Sketch Engine*²⁵ (Kilgarriff et al. 2014), на којој се могу поставити различити упити. Рецимо, можемо извући конкорданце изабране леме или израза, потом колокације за лему, тезаурус сродних речи, затим такозвану скицу речи (енгл. *Word Sketch*) или видети разлику у употреби две сродне речи (енгл. *Word Sketch Difference*). Модул екстракције скице речи, који су развили Kilgarriff et al. (2004), помаже у изградњи Фрејмента и сличних ресурса, убрзава доношење одлука о издвајању различитих значења вишезначних лексема (Baker 2012, 274).

Скица речи даје брзи преглед понашања изабране лексеме тако што прикупља информације из хиљада и милиона примера њене употребе и пружа сажетак категоризованих колокација на једној страници, с везама до појединачних примера. Слика 7 приказује пример скице речи за именицу *ризик* – један поглед на дату страницу довољан је да се

24. JePTex

25. Sketch Engine



Слика 6. Хистограм са фреквенцијама облика обрасца придев-именица за именицу *ризик*

схвати како се та реч користи. Приказан је део скице речи: у првој колони видимо оно што се у српском језику назива предлошко-падежним конструкцијама (у енглеској граматичкој терминологији предлошким фразама):²⁶ *ризик од/са/у/по/на/за...*, за које је даље могуће доћи до конкорданци које одговарају конкретном обрасу кликом на "...". У другој колони представљени су модификатори речи *ризик*: трпни придев глагола (*повећан/идентификован/процењен/...* *ризик*) или придев (*потенцијалан/политички/неприхватљив/...* *ризик*). У трећој колони налазе се глаголи уз које се *ризик* јавља као субјекат, на пример: *смањити се/настајати/постојати/...* Даље следе обрасци у којима се *ризик* јавља као објекат уз глаголе: *смањити/проценити/...*

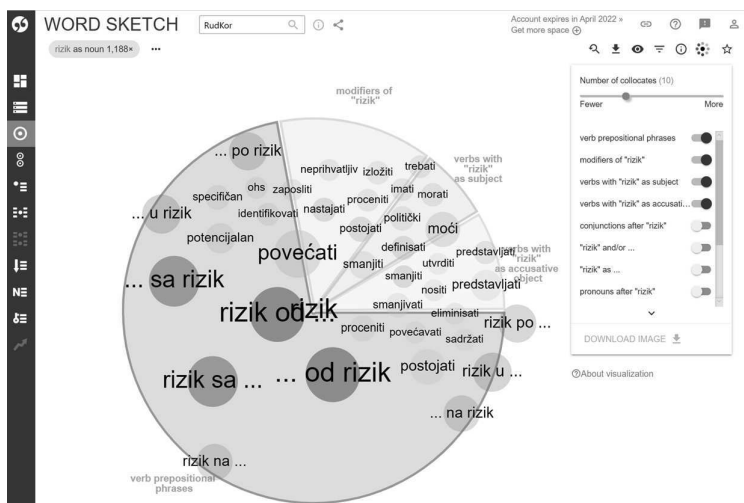
Приказ колокација у виду динамичког дијаграма дат је на слици 8. Може се видети да већину колоката чини образац предлошко-падежне конструкције (на слици: *verb prepositional phrases*). Десни део слике приказује могућности подешавања: који ће обрасци бити приказани и која је минимална фреквенција колокација да би биле укључене на граф.

Истраживање колокација од изузетног је значаја (на пример, у лексикографији, важно је навести најчешће колокате одређених ЛЈ; колокације су кључне и у учењу језика, али и у различитим задацима

26. Важно је напоменути да алати и аутоматско рангирање нису сасвим прилагођени за српски језик и да се јављају грешке, које се морају накнадно исправљати, али су свакако врло драгоцене.

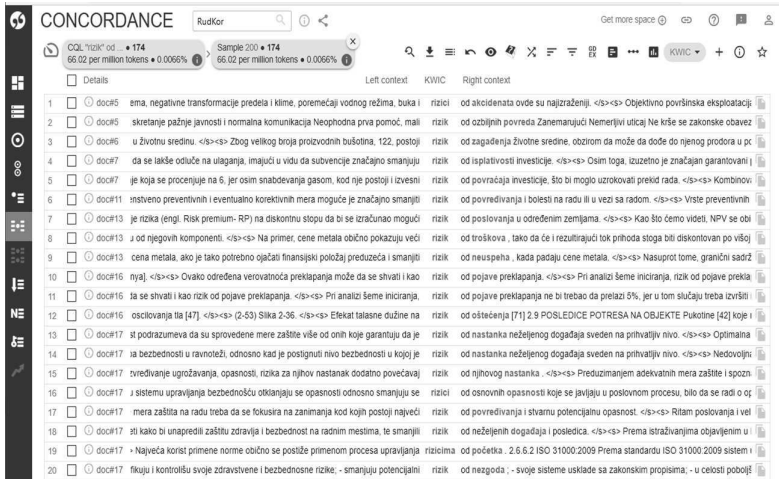
WORD SKETCH				
rizik as noun 1,188				
verb prepositional phrases	modifiers of "rizik"	verbs with "rizik" as subject	verbs with "rizik" as accusative object	conjunctions after "rizik"
"rizik" od ...	povećati	smanjiti	postojati	od
... od "rizik"	radnom mestu sa povećanim rizikom	bi se smanjio rizik	postoji rizik od	sa visokim rizikom od izbijanja požara
"rizik" sa ...	potencijalan	nastajati	smanjiti	kako
... sa "rizik"	potencijalni rizik	preventivnih mera, rizik koji nastaje usled izloženosti zaposlenih	smanji rizik od	rizik kako
... u "rizik"	politički	postojati	proceniti	kao
... u "rizik"	politički rizik	rizik ukoliko postoji	da proceni rizik	mestu sa povećanim rizikom kao i pre premeštaja
"rizik" po ...	identifikovati	definisati	smanjivati	i
... po "rizik"	identifikovani i rangirani rizici	identifikuju opasnosti i rizici, te se definišu potrebne kontrole vezano	u značajnoj meri smanjuje rizik i olakšava proces	sa povećanim rizikom i
"rizik" na ...	proceniti	moći	eliminirati	ili
... na "rizik"	procenjeni rizici	rizik može	način koji bi eliminisao rizik? Korak	mestu sa povećanim rizikom ili za upotrebu
"rizik" za ...	neprihvatljiv	trebati	nositi	da
... za "rizik"	neprihvatljiv Uopšteno - visoki rizik	rizik treba	nose dosta veliki rizik	postoji rizik da
	izložiti	morati	povećavati	
	su zaposleni posebno izloženi rizicima u slučaju nestanka	identifikuju opasnosti i rizici vezane uz promene, organizacija mora osigurati da se	povećava rizik	
	ohs	predstavljati	predstavljati	
	neophodna zbog usvajanja OHS rizicima. To uključuje	utvrditi	predstavlja rizik	
	specifičan	imati	sadržati	
	kajma se pojavljuje specifičan rizik od nastanka povreda	rizik vezan za profitabilnost ima	sadrži minimalni rizik	

Слика 7. Скица речи *ризик* у алату *Sketch engine*



Слика 8. Графички приказ колокација за именицу *ризик* у алату *Sketch engine*

обrade природних језика). Полазећи од скице речи за облик *ризик* и колокацију *ризик од*, могу се приказати одговарајуће конкорданце (слика 9).



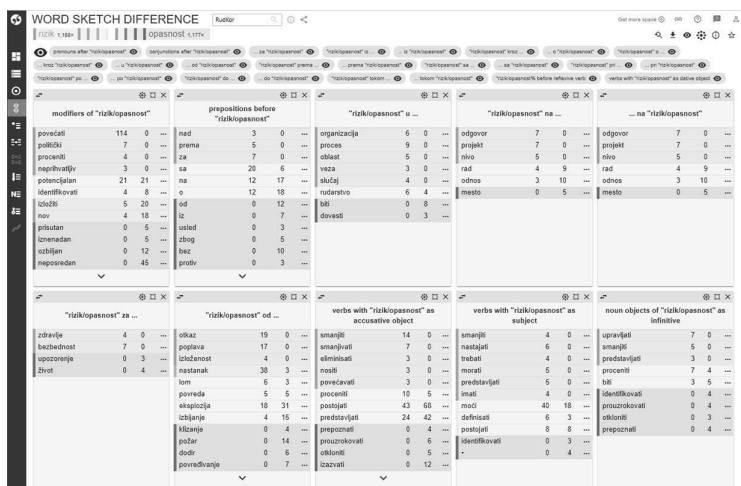
Слика 9. Конкорданце у алату *Sketch engine*

Скица речи даје брзи преглед за унапред припремљена правила, али се корисни резултати могу добити и CQL упитима. На пример, на питање (*Од*) чега је *ризик* (тј. *од чега постоји ризик*)?, одговор бисмо могли наћи следећим CQL упитом [lemma="ризик"][tag="N"]. Одговор на питање *Какав је ризик*? дао би упит [tag="A"][lemma="ризик"]; или, ако желимо да дозволимо размак између речи (не више од 5), онда бисмо написали упит на следећи начин: [tag="V"][word!=""]0,5[lemma="ризик"].

Осим што се може дати испис статистичких фреквенција, оне се могу приказати и визуелно, линијама на десној страни. Слика 10 приказује фреквенције појављивања именице *ризик* уз поједине придеве и глаголе.

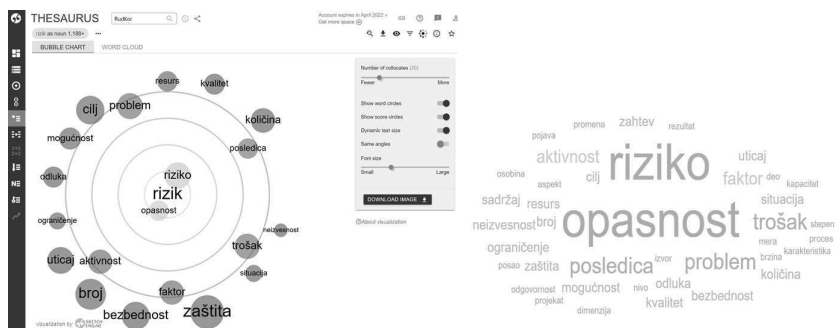
Слика 11 приказује панел разлика у скицама речи (енгл. *Word Sketch Difference*). Панел разлика представља проширење модула скице речи. Генерише скице за две речи и упоређује их, што омогућава да се уоче разлике у њиховој употреби. Функција је посебно корисна за блискозначнице, за антониме и речи из истог семантичког поља. Са

Rows per page: 20 1-20 of 93 < > 1 / 5 > >1



слике 10 се, на пример, може видети да се уз *ризик* користе колокати *повећати*, *политички*, *проценити*, *прихватљив*, док се уз *опасност* јављају *непосредан*, *озбиљан*, *изненадан*.

Аутоматско генерисање тезауруса за задату реч проналази синониме или речи које припадају истој категорији (истом семантичком пољу) и приказује их у виду табеле, с могућношћу позивања одговарајућих конкорданци, скице речи, разлика, тезауруса. Могуће је и графичко приказивање тезауруса, као на слици 12, на којој се могу видети аутоматски генерисане речи које припадају истој категорији (семантичком пољу) као задата реч *ризик*, лево у виду графикона с мехурићима (енгл. *bubble graph*), а десно у виду облака речи. Листа речи тезауруса направљена је на основу контекста у којем се речи појављују у изабраном текстуалном корпусу, ослањајући се на теорију дистрибуционе семантике, која укратко каже да су речи које се појављују у истом контексту сличног значења. Да би се одредили синоними, упоређују се скице речи свих речи исте врсте, а оне које деле највећи део колоката наводе се као сличне речи. Оцена²⁷ дата за сваки синоним указује на проценат заједничких колоката.



Слика 12. Графички приказ тезауруса речи *ризик*

27. Статистичке формуле које се користе у алату *Sketch engine*:
статистика/формуле

5. Закључак

У раду су приказани резултати прелиминарних истраживања у вези с могућностима примене теорије семантике оквира и принципа коришћених у изради семантичке мреже Фрејмнет на примерима из домена оквира ризика прилагођених српском језику. Уз то, представљена је платформа NLTK, погодна за коришћење разних језичких ресурса, као и систем *Sketch Engine* за корпусну анализу текста.

Показано је да анализа у оквиру Фрејмнета нуди детаљну, структурирану интерпретацију, која се даље може користити у различитим апликацијама за обраду природног језика, посебно у проналажењу и смисленом организовању текстова и у природнијој интеракцији између човека и рачунара, нарочито у све популарнијим дијалог-системима (енгл. *chatbot*). Систем дијалога који комуницира с корисницима мора бити у стању да препозна различите речи које призивају исту ситуацију или се односе на исти ентитет да би се извршило успешно препознавање намере.

Погодно је то што се енглески лексикон Фрејмнет може попунити лексичким подацима из других језика, рецимо српског (уз задржавање оних информација о неком оквиру које су заједничке за два језика и модификацију оних које су језички специфичне), тако да је модел погодан за вишејезичне ресурсе и апликације.

Представљена истраживања тек наговештају могућности прилагођавања Фрејмнета за српски језик и његовог поравнавања с таквим базама података за друге језике у будућности. Планирана даља истраживања треба да обухвате и могућности поравнања и заједничког коришћења српског Ворднета и Фрејмнета, као и њихово узајамно допуњавање. У том послу руководићемо се препорукама које дају Tonelli and Pighin (2009).

Овим истраживањима подстиче се и развој корпусне лексикографије српског језика, као и модернизација његовог граматичког и лексикографског описа. Тој модернизацији свакако би допринеле и студије случаја у којима би се поредиле обраде неких вишезначних лексема у Речнику САНУ с њиховим описом из угла теорије семантике оквира.

Могућности за будућа истраживања у овој области су велике. Примена теорије семантике оквира и методологије која се користи у изради Фрејмнета, као и представљених алата на српски језик захтеваће решавање бројних питања; на основу овог рада сматрамо да

ће бити веома изазовно сагледати уз помоћ појма нултог појављивања (*null-instantiation*) допуне транзитивних глагола које се не морају експлицирати, али се подразумевају (*кувати*, *писати* и сл.), као и начин на који се у описним речницима представљају такви случајеви. Уз то, сматрамо да би случајеве нултог појављивања (а у методологији Фрејмнета су препозната три таква случаја) било добро увести у српску граматику.

Захвалност

Овај рад је подржало Министарство просвете, науке и технолошког развоја Републике Србије према уговору бр. 451-03-9/2021-14, који је склопљен са Институтом за српски језик САНУ. Приступ алату SketchEngine је обезбеђен пројектом ELEXIS који се финансира из Horizon 2020 истраживачког и иновационог програма Европске уније, грант број 731015.

Литература

- Atkins, Beryl T. S. 1994. "Analyzing the verbs of seeing: a frame semantics approach to corpus lexicography." In *Annual Meeting of the Berkeley Linguistics Society*, 20:42–56. 1.
- Atkins, Sue, Charles J Fillmore, and Christopher R Johnson. 2003. "Lexicographic relevance: Selecting information from corpus evidence." *International Journal of lexicography* 16 (3): 251–280.
- Baker, Collin F. 2012. "FrameNet, current collaborations and future goals." *Language Resources and Evaluation* 46 (2): 269–286.
- Boas, Hans C., and Ryan Dux. 2017. "From the past into the present: From case frames to semantic frames." *Linguistics Vanguard* 3 (1): 20160003. <https://doi.org/doi:10.1515/lingvan-2016-0003>.
- Brač, Ivana, and Ana Ostroški Anić. 2019. "From concept definitions to semantic role labeling in specialized knowledge resources." In *Proceedings of the 13th International Conference of the Asian Association for Lexicography*, 604–611.

- Fillmore, Charles J. 1976. "Frame semantics and the nature of language." In *Annals of the New York Academy of Sciences: Conference on the origin and development of language and speech*, 280:20–32. 1. New York.
- Fillmore, Charles J. 1982. "Frame semantics." *Linguistic society of Korea (ed.), Linguistics in the morning calm*, 111–137.
- Fillmore, Charles J, and Beryl T Atkins. 1992. "Toward a frame-based lexicon: The semantics of RISK and its neighbors." *Frames, fields and contrasts: New essays in semantic and lexical organization* 75:102.
- Fillmore, Charles J, and Sue Atkins. 1994. "Starting where the Dictionaries Stop: The Challenge of Corpus Lexicography." In *Computational Approaches to the Lexicon*, edited by Sue Atkins and Antonio Zampolli, 349–393. Oxford: OUP.
- Fillmore, Charles J, Miriam RL Petruck, Josef Ruppenhofer, and Abby Wright. 2003. "FrameNet in action: The case of attaching." *International journal of lexicography* 16 (3): 297–332.
- Gantar, Polona, Kristina Štrkalj Despot, Simon Krek, and Nikola Ljubešić. 2018. "Towards semantic role labeling in Slovene and Croatian." In *Proceedings Conference on Language Technologies and Digital Humanities in Ljubljana*, 93–98.
- Gildea, Daniel, and Daniel Jurafsky. 2002. "Automatic labeling of semantic roles." *Computational linguistics* 28 (3): 245–288.
- Hamilton, Craig, Svenja Adolphs, and Brigitte Nerlich. 2007. "The meanings of 'risk': A view from corpus linguistics." *Discourse & Society* 18 (2): 163–181.
- Jurafsky, Dan, and James H Martin. 2020. "Semantic Role Labeling and Argument Structure." Chap. 19 in *Speech and Language Processing*, 3rd ed. December 30, 2020 draft.
- Kilgarrieff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubiček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychl, and Vít Suchomel. 2014. "The Sketch Engine: ten years on." *Lexicography* 1 (1): 7–36.
- Kilgarrieff, Adam, Pavel Rychly, Pavel Smrz, and David Tugwell. 2004. "Itri-04-08 the sketch engine." *Information Technology* 105 (116).

- Kitanović, Olivera. 2021. “Ontološki model upravljanja rizikom u rudarstvu.” PhD diss., Univerzitet u Beogradu, Rudarsko-geološki fakultet. <https://uvidok.rcub.bg.ac.rs/bitstream/handle/123456789/4305/Dokorat.pdf?sequence=1>.
- Kitanović, Olivera, Ranka Stanković, Aleksandra Tomašević, Mihailo Škorić, Ivan Babić, and Ljiljana Kolonja. 2021. “A Data Driven Approach for Raw Material Terminology.” *Applied Sciences* 11 (7): 2892.
- Krstev, Cvetana. 2008. *Processing of Serbian. Automata, texts and electronic dictionaries*. Faculty of Philology of the University of Belgrade.
- Krstev, Cvetana, and Duško Vitas. 2005. “Corpus and Lexicon-Mutual Incompleteness.” In *Proceedings of the Corpus Linguistics Conference*, 14:17.
- Pradhan, Sameer, Wayne Ward, Kadri Hacioglu, James H Martin, and Dan Jurafsky. 2005. “Semantic role labeling using different syntactic views.” In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL’05)*, 581–588.
- Rasulić, Katarina, and Duška Klikovac. 2014. *Jezik i saznanje: hrestomatija iz kognitivne lingvistike*. Univerzitet u Beogradu, Filološki fakultet.
- Ruppenhofer, Josef, Michael Ellsworth, Myriam Schwarzer-Petruck, Christopher R Johnson, and Jan Scheffczyk. 2016. *FrameNet II: Extended theory and practice*. Revised November 1, 2016. https://framenet.icsi.berkeley.edu/fndrupal/the_book.
- Stanković, Ranka, Cvetana Krstev, Biljana Lazić, and Mihailo Škorić. 2018. “Electronic dictionaries-from file system to lemon based lexical database.” In *Proceedings of the 11th LREC - W23 6th Workshop on Linked Data in Linguistics: Towards Linguistic Data Science (LDL-2018), LREC 2018, Miyazaki, Japan, May 7-12, 2018*, 48–56.
- Stanković, Ranka, Branislava Šandrih, Cvetana Krstev, Miloš Utvić, and Mihailo Skoric. 2020. “Machine Learning and Deep Neural Network-Based Lemmatization and Morphosyntactic Tagging for Serbian.” In *Proceedings of The 12th LREC – Language Resources and Evaluation Conference*, 3954–3962.

- Tomašević, Aleksandra, Ranka Stanković, Miloš Utvić, Ivan Obradović, and Božo Kolonja. 2018. "Managing mining project documentation using human language technology." *The Electronic Library*, <https://doi.org/10.1108/EL-11-2017-0239>.
- Tonelli, Sara, and Daniele Pighin. 2009. "New features for framenet-wordnet mapping." In *Proceedings of the thirteenth conference on computational natural language learning (CoNLL-2009)*, 219–227.
- Utvić, Miloš. 2011. "Annotating the Corpus of Contemporary Serbian." *IN-FOtheca* 12, no. 2 (December): 36a–47a.
- Wasserscheidt, Philipp, and Andrea Hrstić. 2020. "Legal Variation? A Frame Analysis of Croatian and Serbian in the Domain of Law." *Mediterranean Language Review* 27:87–112.
- Драгићевић, Рајна. 2007. *Лексикологија српског језика*. Београд: Завод за уџбенике.
- Марковић, Александра. 2017. "Однос граматике и речника – граматика инхерентна описним речницима српског језика." *Наш језик XLVIII* (1-2): 27–43.
- Поповић, Љубомир. 2003. "Интегрални речнички модели и њихов значај за лингвистички опис и анализу корпуса." *Научни састанак слависта у Вукове дане* 31 (1): 201–220.

О Корпусу студената англистике (КорСАНг) и могућностима његове софтверске експлоатације

УДК 81'322.2

САЖЕТАК: Корпусна лингвистика у Босни и Херцеговини и Републици Србији не користи се довољно. Разлози за то су многи: недостатак језичких корпуса, бојазан од рачунарских метода, још увијек присутан традиционални приступ обраде података који је квалитативан или не сеже даље од дескриптивне статистике. Неријетко ћемо чути аргументе против рачунарског метода, као што је тај да велики број примјера који су резултат претраге може нарушити квалитет анализе и довести до извођења погрешних лингвистичких законитости. Ипак, развој технологије утицао је и на развој информатичке писмености у свим људским дјелатностима, укључујући и академску заједницу. У овом раду образложићемо на који начин је неколико катедри за англистику у Босни и Херцеговини и Србије у сарадњи са Друштвом за језичке технологије и ресурсе понудило једно рјешење ученичког корпуса, те описати како је текао процес прикупљања корпуса и какве опције претраге корпус нуди.

КЉУЧНЕ РЕЧИ: ученички корпус, корпусна лингвистика, паралелни корпуси, анотација корпуса

РАД ПРИМЉЕН: 16. август 2021.

РАД ПРИХВАЋЕН: 2. септембар 2021.

Миња С. Радоња

minja.radonja@ff.ues.rs.ba

Срђан Р. Шућур

srdjan.sucur@ff.ues.rs.ba

Универзитет у

Источном Сарајеву

Филозофски факултет Пале

Сарајево

Босна и Херцеговина

1. Увод

„Компјутерски ученички корпуси електронске су базе података формиране од аутентичне дискурсне продукције на страном/другом

језику по експлицитним критеријумима, тако да се могу користити за истраживање у оквиру усвајања другог језика и/или наставе страног језика. Базе података бележе се стандардизовано и хомогено, уз информације о пореклу података“ ((Granger 2002, 7), према (Марковић 2019, 13)). Корпус студената англистике (КорСАНг) ученички је корпус студената англистике србофоних говорника са 4 универзитета – Универзитет у Источном Сарајеву, Универзитет у Бањалуци, Универзитет у Београду и Универзитет у Новом Саду. Корпус је настао као производ два значајна национална пројекта: *Фразеолошка компетенција српских говорника енглеског кроз призму контрастивне анализе међујезика*, 19/6-020/961-46/18, који је реализован у периоду од 31. 12. 2018, и *Научни потенцијали аотираних ученичких корпуса у примијењеној лингвистици*, 19.032/961-135/19, који је реализован у периоду од 31. 12. 2019. Корпус КорСАНг је инспирисан првим ученичким корпусом енглеског језика србофоних говорника *ICLE-SE*, који је формиран у периоду од септембра 2015. до септембра 2016. године и састоји се од аргументативних састава студената англистике србофоних говорника писаним на енглеском језику. Досадашњи резултати истраживања на корпусу *ICLE-SE* представљени су у научним радовима (Марковић 2017, 2018, 2019, 2020; Radić-Bojanić 2019; Radonja 2019; Šućur 2019; Spajić and Suknović 2019; Tomović and Stefanović 2019). Више о корпусу *ICLE-SE* и начинима његове претраге видјети у (Шућур 2020).

1.1 Структура корпуса

КорСАНг обухвата 327 текстова и укупно има око 140.000 речи. По саставу је трокомпонентан, тј. састоји се од три поткорпуса:

- Поткорпус превода студената англистике са енглеског језика (*КорПСАНг1*), који се састоји од превода књижевних (3 текста) и новинских чланака (3 текста) са енглеског на српски језик. Овај дио корпуса броји 127 задатака превода и 46.785 ријечи. То практично значи да се сваки превод може упарити са оригиналним текстом који је у овом поткорпусу на енглеском језику, а што се назива битекстом.^{1.}

1. „[Б]итекст представља текст и његов превод, односно преводе, представљене на такав начин да је између елемената њиховог логичког исказа успостављена експлицитна веза, на пример, на нивоу пасуса или реченица“ ((Vitas 2010, 273), према (Андоновски 2019, 17))

- Поткорпус превода студената англистике на енглески језик (*КорПСАн2*), гдје су почетни текстови из жанра књижевности (1 текст) и новинског жанра (5 текстова). Коначан број превода овог поткорпуса је 130, а број ријечи 50.128. Као и у претходном поткорпусу, и *КорПСАн2* је паралелни корпус.
- Поткорпус аргументативних састава² студената англистике писаних на српском језику као матерњем (*КорССАн2*) обухвата 70 састава и 40.012 ријечи.

Прве двије компоненте јесу паралелни ученички корпуси, док је трећа компонента осмишљена као референтни корпус корпусу *ICLE-SE*. „Под паралелним корпусима подразумевају се корпуси који садрже један или више оригиналних текстова и њихове преводе на један или више језика“ ((Топу 2016, 11) према (Андоновски 2019, 16)). О значају и примјени паралелних корпуса видјети детаљније (Андоновски 2019, 2021; Ристовић 2012, 2016)).

У наставку рада, у другом одјељку ћемо описати фазе припреме корпуса, од прикупљања и одабира текстова, преко обраде и паралелизације текстова, до израде метаподатака и формирања корпуса. Трећи одјељак приказује како се корпус може претраживати. Коначно, у закључку, сумирамо досадашње резултате и представљамо даље планове.

2. Припрема корпуса *КорСАн2*

У формирању корпуса укупно је учествовало 194 студената англистике. У периоду од 2016. до 2019. године формиран је *КорССАн2*, а од 2017. до 2019. године *КорПСАн1* и *КорПСАн2*. Од наведене три компоненте, најзахтјевније је било прикупити корпус састава на српском језику, због недовољне жеље студената да учествују у писању састава на матерњем језику, те је формирање овог дијела корпуса најдуже потрајало. Пројекат настанка паралелног *Корпуса студената англистике* састојао се од неколико етапа:

- припремна фаза,

2. „Аргументативни текст (расправа) је текст у којем аутор почиње од неке сумње или дилеме, а онда навођењем аргумената долази до разрешења. Аутор таквог текста тежи да докаже саговорника одн. читаоца убеди у исправност свог мишљења“ (СЕО 2015, 79).

- прикупљање текстова,
- обрада текстова,
- паралелизација,
- вођење евиденције о метаподацима,
- публикување текстова и метаподатака на платформу.

2.1 Припремна фаза

У припремној фази је дефинисано неколико кључних питања у вези са пројектом: главни и споредни циљеви пројекта; одређивање обима пројекта; прецизирање сарадње и расподјела задатака између чланова тима са Универзитета у Источном Сарајеву, Универзитета у Бањалуци, Универзитета у Београду и Универзитета у Новом Саду; прецизирање сарадње са члановима тима из ЈеРТеха;³ утврђивање временског оквира извршења појединачних етапа. За успостављање контаката са члановима тима била је задужена руководица пројекта Јелена Марковић.

2.2 Прикупљање и одабир текстова

У сљедећој фази су чланови тима, заједно са руководиоцем пројекта, извршили одабир неколико текстова за превод са енглеског и на енглески језик и направили прелиминарни списак тема састава на српском језику. Приликом селекције текстова, водило се рачуна да текстови буду из књижевног и новинског жанра. Новински чланци покривају различите теме, као што су политика, економија, језик, здравље, друштвене мреже, шоу бизнис. У коначници је изабрано 6 текстова на енглеском, 6 текстова на српском и 24 теме састава на српском језику. Потом се приступило сакупљању превода и састава, што је био временски најзахтјевнији задатак који се састојао од организовања услова за преводње текстова и писање састава од стране студената. Чланови тима су организовали овај задатак на својим факултетима. Преводње је било временски ограничено, са изузетком од неколико радова, док је писање састава бар у трећини радова било неограничено. Уз задате писмене задатке, студенти су били дужни попунити пропратни документ профила учесника како би дали сагласност да се њихови преводи и есеји користе у сврхе истраживања, док су сами радови анонимни и евидентирани под шифром.

3. Друштво за језичке технологије и ресурсе – ЈеРТех

2.3 Обрада и паралелизација текстова

Радни задаци који се тичу обраде текстова и паралелизације обављени су од стране сарадника на пројекту, под руководством Јелене Марковић и тима из ЈеРТеха. Сарадници су прошли онлајн обуку за коришћење алата за паралелизацију како би се добили текстови у формату TMX⁴ (Translation Memory eXchange) докумената. Упознати су са радом *Notepad++*⁵ XML едитора, *Unitex/GramLab*,⁶ *Unitex* модула за српски (Krstev 2008) и интегрисаног развојног окружења за паралелизоване корпусе *ACIDE* (Aligned Corpora Integrated Development Environment) (Utvić, Stanković, and Obradović 2008), који су омогућили адекватну обраду текста по прописаним правилима. У *Notepad++* извршена је припрема изворног и преведеног текста, која подразумева упаривање параграфа два документа, тако да сваком параграфу изворног текста одговара параграф превода. Тако припремљени фајлови даље су обрађени у *Unitex Visual IDE*, који омогућава сегментацију параграфа на реченице, сагласно језику на ком је писан текст. Након сегментације оригиналног и преведеног текста, фајлови су припремљени за паралелизацију, тј. „процес успостављања веза између одговарајућих варијанти јединица превођења, односно формирање сета јединица превођења“ (Андоновски 2019, 17). Паралелизација је омогућена коришћењем апликације *ACIDE* коју је развила Група за језичке технологије Универзитета у Београду (више о интегрисаном развојном окружењу за паралелизоване корпусе *ACIDE* може се пронаћи код (Utvić, Stanković, and Obradović 2008)). Ова апликација омогућава аутоматску паралелизацију уз могућност провјере и корекција погрешака, што је члановима тима који су имали овај задатак помогло да успјешно припреме текстове за паралелни корпус. У текстовима који су припремани за паралелни корпус у оквиру овог

4. TMX је ISO стандард (ISO24616 2012) за складиштење такозваних преводилачких меморија (Translation Memories) и њихову размену између различитих софтверских преводилачких алата, као и између различитих фирми које се баве одржавањем преводилачких меморија (TMX 2005). Преводилачке меморије представљају збирке одредница у којима је текст изворног језика повезан са еквивалентним преводом текста циљног језика односно произведени TMX документ састављен је од добијених јединица превођења (Андоновски 2019, 22).

5. *Notepad++*

6. *Unitex/Gramlab*, Пакет за обраду корпуса применљив на различитим платформама.

пројекта, наилазили смо на примјере да у тексту превода недостају неки сегменти, или једном сегменту у оригиналном тексту одговара неколико сегмената у тексту превода, што се може јавити код превода. Сама обука је била изазов тиму лингвиста који се раније нису сусретали са оваквом врстом радних задатака, а стечене вјештине у току израде паралелног корпуса даље су нову перспективу на корпусну лингвистику. Као коначни резултат након обраде текста у *ACIDE*-у је *TMX* документ који се уноси у корпус.

2.4 Израда метаподатака и формирање корпуса

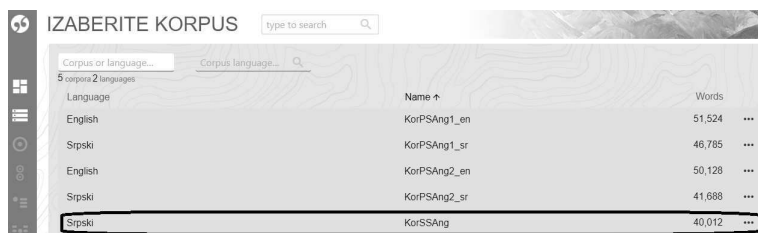
Паралелно са припремом текстова за корпус припремани су метаподаци који садрже информације о варијаблама као што су подаци о учеснику (узраст, пол, матерњи језик, подаци о образовању, године учења енглеског, познавању других страних језика, боравак на енглеском говорном подручју) и варијабле контекста језичке продукције, тј. подаци о томе у каквом су контексту ученици извршавали задатке (да ли је постојало временско ограничење за задатке писања састава или превода, коришћење рјечника и приручника, да ли је задатак испитна обавеза или је писан ван испита, који је жанр, и сл.). Табеле са метаподацима рађене су посебно за преводе и саставе у засебним Excel документима. С обзиром да су радови који улазе у корпус анонимни, сваком документу је било потребно додијелити одговарајуће метаподатке заведене под истом шифром. Чланови тима из ЈеРТеха припремљене документе су припремили за публикување: додијелили метаподатке сваком документу, извршили анотацију врстом ријечи и лематизацију (детаљније о ресурсима који су омогућили ово видјети у (Stanković et al. 2020)).

3. Претрага корпуса *КорСАНг*

Укратко смо описали како је текао процес прикупљања корпуса од припремне фазе све до његовог коначног електронског облика и изазови са којима смо се на том путу сусретали. *КорСАНг* садржи 136.925 ријечи: *КорПСАнг1* (46.785 ријечи, 127 превода на српски језик), *КорПСАнг2* (50.128 ријечи, 130 превода на енглески), *КорССАНг* (40.012 ријечи, 70 састава на српском језику). Један од већих изазова је био мотивисати студенте да пишу саставе на матерњем језику, што буди сумњу у опадање компетенције писања на матерњем језику (Šućur 2020, 144).

Напоменућемо и то да теме састава ограничавају могућност претраге. Од 24 понуђене теме, више од 40% састава је писано на четири теме: *Породица или посао: шта је важније*, *Како музика утиче на живот*, *Како замишљам доброг родитеља*, и *Од склапања брака до развода данас*. Стога очекујемо да корпус састава нуди богат вокабулар који припада семантичком пољу породичних односа, али сиромашнији када су у питању неке друге понуђене теме. У наставку ћемо, након уводних напомена о платформи *Sketch Engine*, најприје показати како је помоћу ње могуће вршити основну и напредну претрагу корпуса КорССАнг, а затим и претраживање паралелизованих корпуса *КорПСАнг1_ен* (чине га изворни текстови на енглеском) и *КорПСАнг1_ср* (чине га преводи на српски), односно *КорПСАнг2_ср* (чине га изворни текстови на српском) и *КорПСАнг2_ен* (чине га преводи на енглески).

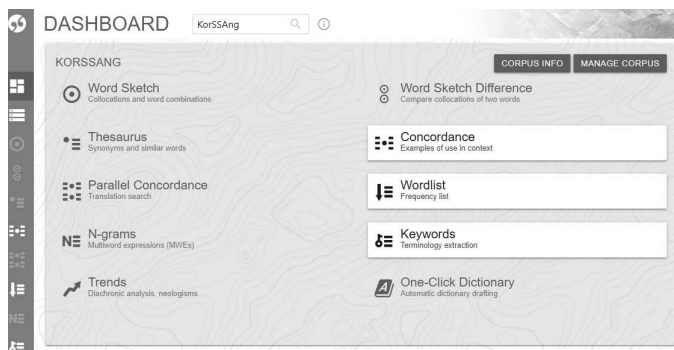
3.1 *Sketch Engine*



Слика 1. Одабир корпуса

Sketch Engine је онлајн платформа сачињена од низа робустних алата за претраживање електронских корпуса. Ову платформу је, у сарадњи са чешким програмером Павелом Рихлијем [*Pavel Rychlý*], осмислио британски лексикограф и корпусни лингвиста Адам Килгариф [*Adam Kilgariff*], а њена комерцијална верзија доступна је од 2004. године (Kunilovskaya and Koviiazina 2017, 503). Како истиче Kilgariff et al. (2014, 15-16), *Sketch Engine* је потекао из свијета академских истраживања, гдје се данас користи у лингвистици и на катедрама за изучавање језика (у настави и истраживањима), затим у рачунаској лингвистици (енг. *Computational Linguistics*) и обради природног језика (енг. *Natural Language Processing*). Поред тога, користи се и у настави језика

(нарочито енглеског као страног), лексикографији, превођењу и настави превођења, затим анализи дискурса итд. (Kilgariff 2014, 15-16).



Слика 2. Контролна табла

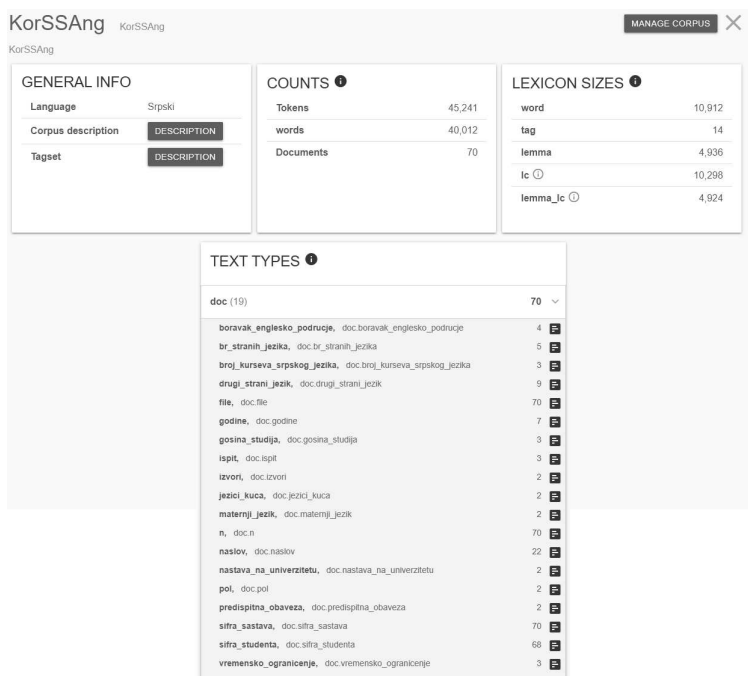
У наставку ћемо представити широко доступну некомерцијалну верзију ове платформе коју је прилагодио ЈеРТех, која и поред ограниченог броја алата нуди мноштво могућности за претраживање и изучавање (паралелизованих) електронских корпуса ученичке продукције на матерњем и енглеском језику.

3.2 Пример претраге поткорпуса *КорССАнг*

У првом кораку (слика 1) бирамо један од пет доступних корпуса, *КорССАнг*, кога чине аргументативни састави на српском језику чији су аутори србофони ученици енглеског.

Тиме приступамо контролној табли (слика 2) која нуди неколико алата, и то: конкорданце (енг. *Concordance*), чијим одабиром је могуће вршити претрагу употребе неког термина у одговарајућем контексту, затим листе ријечи (енг. *Wordlist*), којим се термини претражују према учесталости, као и трећи алат – кључне ријечи (енг. *Keywords*) који може бити од користи у лексикографским истраживањима, при издвајању термина и сл.

У овом корисничком сучељу могуће је провјерити и метаподатке о одабраном корпусу (енг. *Corpus Info*) (слика 3).



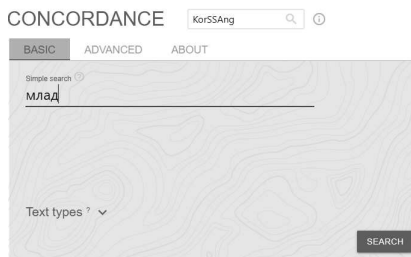
Слика 3. Метаподаци о корпусу *KorSSAng*

Овдје су доступни општи подаци о корпусу који садрже његов опис,⁷ скуп ознака којима је етикетиран,⁸ те број токена, ријечи, лема и докумената које га чине, а одабиром опције Manage corpus (горњи десни угао) могуће је модификовати корпус у смислу проширивања, нпр. додавањем елемената постојећем корпусу (енг. *Make bigger* → *Add texts*

7. KorSSang

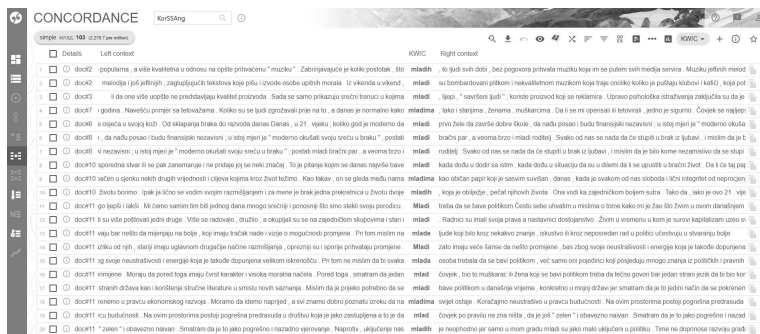
8. 1. N (именица, енг. *Noun*), 2. A (придјев, енг. *Adjective*), 3. V (глагол, енг. *Verb*), 4. PRO (замјеница, енг. *Pronoun*), 5. NUM (број, енг. *Number*), 6. PREP (приједлог, енг. *Preposition*), 7. CONJ (везник, енг. *Conjunction*), 8. INT (узвик, енг. *Interjection*), 9. PAR (партикула, енг. *Particle*), 10. ADV (прилог, енг. *Adverb*), 11. PREF (префикс, енг. *Prefix*), 12. ABB (скраћеница, енг. *Abbreviation*), 13. RN (римски број, енг. *Roman numeral*), 14. PUNCT (интерпункцијски знак, енг. *Punctuation*), 15. SENT (ознака за крај реченице, енг. *Sentence end marker*), 16. ? (ријечи које нису српског поријекла, или суфикси у сложеницама, енг. *Non-Serbian words or suffixes in compounds*).

to corpus), или формирањем и измјеном поткорпуса (енг. *Subcorpora* → *Manage subcorpora*) и сл.



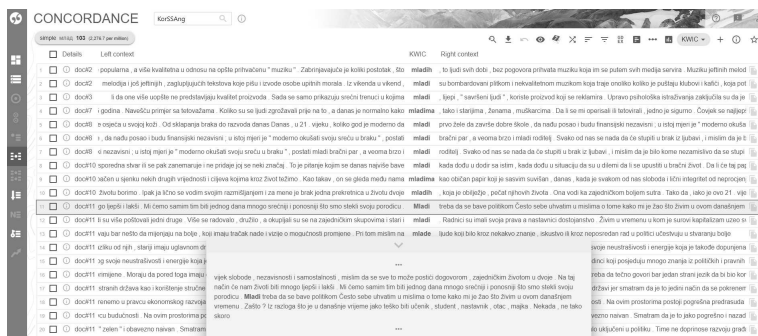
Слика 4. Основна претрага леме „млад“

Даље, одабиром менија за претрагу конкорданци, у основном типу претраге (енг. *Basic*), користећи ћирилично писмо, вршимо упит за лему „млад“ (слика 4), пошто су аутори састава припадници млађе популације, па се на датом корпусу може очекивати значајна дистрибуција овог термина.



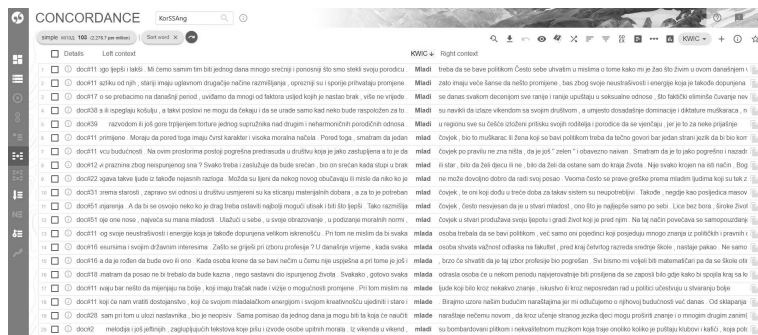
Слика 5. Резултати основне претраге за лему „млад“

Претрагом добијамо 103 резултата које чине облици придијева „млад“, као и облици поимениченог придијева „млади“, поређаних редослиједом којим се јављају у нумерисаним документима који чине поткорпус КорССАнг (слика 5).



Слика 6. Шири контекст употребе одабране леме „млад“

Одабиром неког од издвојених резултата добија се шири контекст у коме се дата лема јавља (слика 6).



Слика 7. KWIC леме „млад“.

Добијене резултат претраге могуће је поређати према лијевом и десном контексту, или према кључним ријечима у контексту (енг. *Key Word in Context - KWIC*), који је и најубочијанији формат за приказивање низова конкорданци (слика 7), те их је могуће сачувати у неколико формата, и то: *TXT*, *CSV*, *XLS*, *XML*, а тренутни приказ могуће је сачувати и као *PDF*.

Платформа омогућава проверу учесталости (слика 8) и дистрибуције тражене леме у читавом корпусу, као и термина са којима се најчешће јавља у колокацији (слика 9) на растојању од 1-5 мјеста са лијеве и

Lemma	↓ Frequency	Frequency per million
1 млад	103	2,276.70

Слика 8. Учесталост леме „млад“ у корпусу *КорССАнг*

десне стране тражене леме (у овом случају је то именица људи,⁹ у 25 примјера).

CONCORDANCE *KorSSAng*

Collocations **CHANGE CRITERIA** **BACK TO CONCORDANCE**

Word	Cooccurrences ¹	Candidates ²	T-score	MI	+ LogPice
1 judge	10	39	3.13	6.82	11.17 ...
2 just	15	162	3.78	5.35	10.96 ...
3 people	6	37	2.42	6.15	10.86 ...
4 people	5	28	2.21	6.29	10.29 ...
5 by	5	81	2.15	4.78	9.80 ...
6 media	3	8	1.72	7.36	9.79 ...
7 swaka	3	12	1.72	6.78	9.74 ...
8 ask	3	13	1.71	6.66	9.73 ...
9 told	3	13	1.71	6.66	9.73 ...
10 dates	4	66	1.62	4.73	9.60 ...
11 da	8	250	2.63	3.81	9.54 ...
12 for	7	208	2.47	3.89	9.53 ...
13 could	3	33	1.69	5.32	9.50 ...
14 to	8	262	2.62	3.75	9.49 ...
15 to	5	139	2.09	3.88	9.40 ...
16 to	6	225	2.23	3.49	9.38 ...
17 he	3	69	1.64	4.26	9.36 ...
18	34	1,651	5.11	3.01	9.30 ...
19 he	13	747	3.13	2.83	8.97 ...
20 it	6	297	2.17	3.15	8.94 ...

Rows per page: 20 1-20 of 30 1 2 3 4

Слика 9. Термини са којима се лема „млад“ најчешће јавља у колокацији у корпусу *КорССАнг*

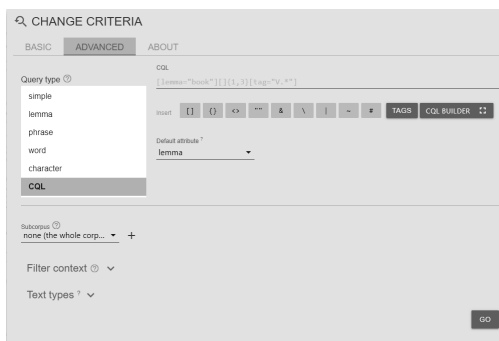
У наставку вршимо напредну претрагу¹⁰ помоћу *CQL* упита (енг. *Context Query Language*), а приступамо јој преко менија приказаног на слици 4, одабиром опције *Advanced* (слика 10).

За потребе овог приказа користимо алат *CQL Builder*, интегрисан на платформи *Sketch Engine* (слика 11), чија сврха је генерисање упита специфичне синтаксе за напредно претраживање корпуса, помоћу скупа ознака из фусноте 8. У нашем случају, тражићемо глаголе који се јављају на растојању од 3 мјеста десно од леме „млад“.

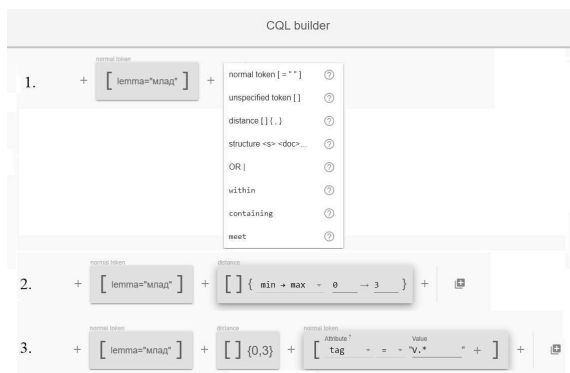
Претрага по овом упиту даје 94 резултата (слика 12), али је неопходно мануелно издвојити јединствене примјере, како би се искључили они који

9. Поређења ради, претрага синтагме „млади људи“ (енг. *young people*) у упоредивом корпусу *LOCNESS* (тј. његовој америчкој компоненти, коју чини приближно 150 хиљада ријечи) даје 12 резултата.

10. Поред претраге *CQL* упитима, напредну претрагу могуће је вршити и помоћу „једноставног“ облика ријечи (енг. *simple*), гдје се претражују сви облици те ријечи који не садрже велико (почетно) слово, затим претрага лема, фраза, ријечи (гдје је свеједно садрже ли велика слова или не), те претрага посебних карактера, попут интерпункцијских знака, бројева, итд.



Слика 10. Одабир типа напредне претраге

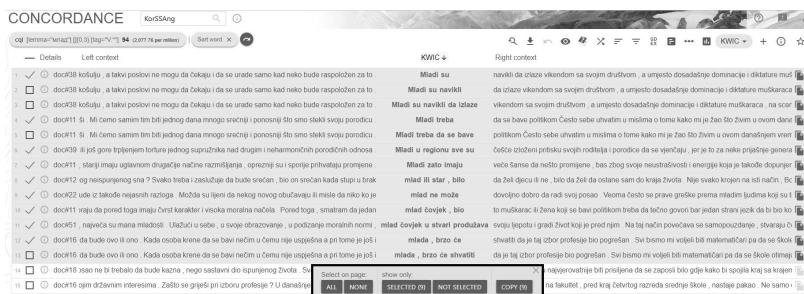


Слика 11. Креирање CQL упита [lemma="млад"] []0,3[tag="V.*"]

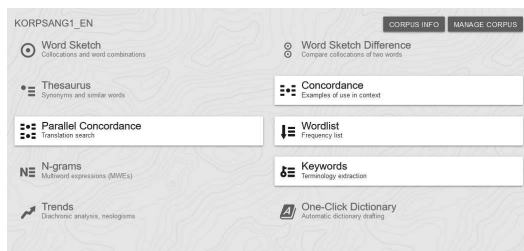
се, због одабира специфичног растојања од 3 мјеста на коме се глагол налази, понављају неколико пута.

У наставу ћемо представити претрагу паралелизованих корпуса *КорПСАНг1_ен* и *КорПСАНг1_ср*, односно *КорПСАНг2_ен* и *КорПСАНг2_ср*. Одабиру ових корпуса такође приступамо преко менија на слици 1.

Најприје бирамо корпус *КорПСАНг1_ен* и приступамо контролној табли (слика 13), која се, самим тиме што је овај корпус паралелизован са корпусом *КорПСАНг1_ср*, од контролне табле на слици 2 разликује по томе што је доступан и алат за претраживање паралелизованих конкорданци (енг. *Parallel Concordance*).



Слика 12. Резултати претраге по CQL упиту [lemma="млад"] [10,3[tag="V.*"]]



Слика 13. Контролна табла за паралелизовани корпус

Затим бирамо алат за претраживање паралелизованих конкорданци и вршимо основну претрагу именичке синтагме *project's collapse* (која је дио веће релативне клаузе – *which then led to the project's collapse*),¹¹ присутне у једном од изворних текстова на енглеском језику, како бисмо проверили на који начин је преведена на српски језик (слика 14).

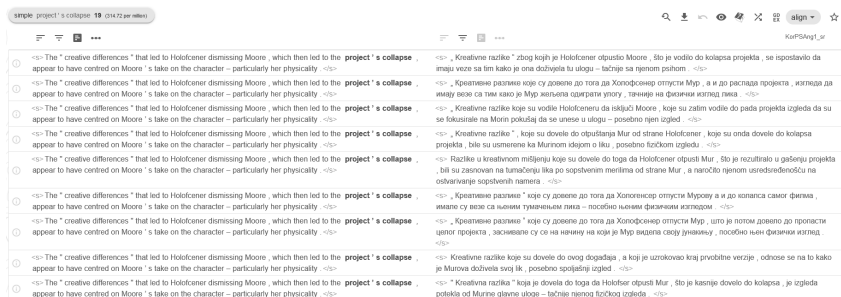
Претрага синтагме *project's collapse* у паралелизованом корпусу даје 19 резултата (слика 15).

У детаљном прегледу добијених резултата може се видјети да је тражена синтагма на енглески језик преведена на неколико различитих начина, претежно именичким синтагмама: *колапс пројекта* (4x), *пропаст (целог) пројекта* (4x), *распад пројекта* (3x), *пад пројекта*, *гашење пројекта*, *колапс (самог) филма*, *колапс*, *крај првобитне верзије*, *крах пројекта*, *пропадање пројекта*, те се, у једном случају, у преводу

11. Напредну претрагу ове синтагме могуће је извршити помоћу сљедећег CQL упита: [lemma="project"] [word=""] [word="s"] [lemma="collapse"].



Слика 14. Основна претрага синтагме *project's collapse*



Слика 15. Резултати основне претраге синтагме *project's collapse*

налази у оквиру клаузе, које су довеле до тога [...] да цео пројекат пропадне.

Како сложени глаголски облици неријетко представљају изазов у учењу страних језика, у наставку вршимо напредну претрагу перфекатских облика, и то помоћу *CQL* упита¹² [lemma="have"] [tag="V.*"] (слика 16).

Претрага по овом упиту даје 432 резултата (слика 17), међу којима су облици прошлог перфекта, садашњег перфекта, перфекатског инфинитива и партиципа.

12. За потребе овог *CQL* упита користимо ознаке којима су етикетиране компоненте корпуса на енглеском, Penn Treebank Tagset.

4. Закључак

У овом раду смо укратко описали кораке настанка Корпуса студената англистике (*КорСАНг*), и представили могућности његове претраге. Познато је да је недостатак нарочито ученичких корпуса хендикеп за истраживаче који се баве примјењеном лингвистиком, те се надамо да ће наши напори да тај проблем пренебрегнемо резултирати већој популарности корпусне лингвистике као истраживачког метода на овим просторима нарочито из области англистичке лингвистике. Досадашњи резултати коришћења корпуса *КорСАНг* су представљени у радовима: (Šušur 2020; Spajić and Suknović 2019; Tomović and Stefanović 2019; Марковић and Станковић, у штампи), а у изради је и једна докторска дисертација чији је дио корпуса базиран на *КорСАНгу*. Научни потенцијал овог корпуса може се очекивати у мноштву научних и стручних радова, монографија и научно-истраживачких пројеката. Додаћемо и то да је у плану израда друге верзије корпуса *КорСАНг*, са проширеним бројем састава и превода, и са могућношћу интегрисања нових софтверских алата који ће пружити лакшу претрагу и угоднију употребу корпуса корисницима.

Захвалност

Овај рад је заснован на истраживањима која су спроведена у оквиру два национална пројекта: *Научни потенцијали аотираних ученичких корпуса у примијењеној лингвистици*, 19.032/961-135/19, и *Фразеолошка компетенција српских говорника енглеског кроз призму контрастивне анализе међујезика*, 19/6-020/961-46/18. Пројекти су суфинансирани од стране *Министарства за научнотехнолошки развој, високо образовање и информационо друштво*, Бања Лука. Захваљујемо *Друштву за језичке ресурсе и технологије – ЈеРТех*, а нарочито проф. Ранки Станковић, на сарадњи на пројекту 19.032/961-135/19. Посебно захваљујемо координатору пројекта, проф. Јелени Марковић, што нам је пружила прилику да учествујемо у поменутих пројектима, као и за сталну подршку у виду охрабрења, савјета, мотивације и стручности.

Литература

СЕО. 2015. *Општи стандарди постигнућа за крај општег средњег и средњег стручног образовања и васпитања у делу општеобразовних*

- предмета. Београд: Завод за вредновање квалитета образовања и васпитања.
- Granger, Sylviane. 2002. "A bird's-eye view of learner corpus research." *Computer learner corpora, second language acquisition and foreign language teaching* 6:3–33.
- ISO24616. 2012. *Language resources management — Multilingual information framework*. International Standard Organization.
- Kilgarrieff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubiček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychl, and Vít Suchomel. 2014. "The Sketch Engine: ten years on." *Lexicography* 1 (1): 7–36.
- Krstev, Cvetana. 2008. *Processing of Serbian. Automata, texts and electronic dictionaries*. University of Belgrade, Faculty of Philology.
- Kunilovskaya, Maria, and Marina Koviagina. 2017. "Sketch engine: A toolbox for linguistic discovery." *Jazykovedny Casopis* 68 (3): 503.
- Radić-Bojanić, Biljana B. 2019. "NEODREĐENA ZAMENICA ONE U PISANJU KOD NEIZVORNIH GOVORNIKA ENGLESKOG JEZIKA." *Годишњак Филозофског факултета у Новом Саду* 44 (2): 39–52. <https://doi.org/10.19090/gff.2019.2.39-52>.
- Radonja, Minja S. 2019. "The use of interactive metadiscourse in Serbian students." *Радови Филозофског факултета: Часопис за хуманистичке и друштвене науке* 8 (21). <https://doi.org/10.7251/FIN1921121R>.
- Spajić, Sonja, and Mina Suknović. 2019. "The Choice of Lexemes According to Their Frequency in Translation into L2." *Komunikacija i kultura online* 10 (10): 104–119. <https://doi.org/10.18485/kkonline.2019.10.10.6>.
- Stanković, Ranka, Branislava Šandrih, Cvetana Krstev, Miloš Utvić, and Mihailo Skoric. 2020. "Machine Learning and Deep Neural Network-Based Lemmatization and Morphosyntactic Tagging for Serbian." In *Proceedings of The 12th Language Resources and Evaluation Conference*, 3954–3962.
- Šućur, Srđan. 2019. "Distribucija frazalnih glagola u pisanju na engleskom kao stranom kod srbofonih govornika." *Komunikacija i kultura online* 10 (10): 120–143. <https://doi.org/10.18485/kkonline.2019.10.10.7>.

- Šućur, Srđan. 2020. "REVERSE TRANSFER IN ADULT SERBIAN EFL LEARNERS' WRITING 1 2-A CORPUS BASED STUDY." *BEYOND HERMENEUTICS*, 141.
- TMX. 2005. "Translation Memory eXchange (TMX) 1.4b Specification." Accessed 1.08.2021. <https://www.gala-global.org/knowledge-center/industry-development/standards/lisa-oscar-standards>.
- Tomović, Nenad, and Sofija Stefanović. 2019. "Uticaj L2 i leksički i leksičko-sintaksički kalkovi u prevodu. Studija slučaja." *Komunikacija i kultura online* 10 (10): 144–154. <https://doi.org/kkonline.2019.10.10.8>.
- Töny, Luzius. 2016. "Corpora als Ressourcen für die maschinelle Übersetzung." Accessed 17.04.2016. https://swanrad.ch/downloads/mt_1.pdf.
- Utvić, Miloš, Ranka Stanković, and Ivan Obradović. 2008. "Integrisano okruženje za pripremu paralelizovanog korpusa." *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen*, 563–578.
- Vitas, Duško. 2010. "Resursi i metode za obradu srpskog – stanje i perspetive." In *Srpska lingvistika/Serbische Linguistik, Eine Bestandsaufnahme, Studies of Language and Culture in Central and Eastern Europe (SLCCEE*, edited by Biljana Golubović and Cristian Voß, 7:257–277. München: Verlag Otto Sagner.
- Андоновски, Јелена. 2019. "Мрежа отворених података и језички ресурси у процесу изградње српско-немачког литерарног корпуса." PhD diss., Универзитет у Београду, Филолошки Факултет, January.
- Андоновски, Јелена. 2021. "Паралелни корпуси у Србији — могућности за паралелно проналажење информација на два или више језика" [in serbian]. 3, *Библиотекар* 63 (1): 51–74. ISSN: 0006-1816. <https://doi.org/10.18485/bibliotekar.2021.63.1.3>.
- Марковић, Јелена. 2017. "Лични метадискурс у писању код изворних и неизворних говорника енглеског језика." *Филолог-часопис за језик, књижевност и културу*, no. 15, 44–60. <https://doi.org/10.21618/fil1715044m>.

- Марковић, Јелена. 2018. “Употребе глагола *маке* у писању на енглеском језику као страном код изворних говорника српског језика (корпуснолингвистичка анализа).” *Зборник матице српске за филологију и лингвистику* 61 (1): 165–180. https://www.maticasrpska.org.rs/stariSajt/casopisi/ZMSFL_61_1.pdf.
- Марковић, Јелена. 2019. *Кроз призму контрастивне анализе међујезика*. Филозофски факултет.
- Марковић, Јелена. 2020. “Концесивни конектори *though* и *however* у писању на енглеском језику код изворних и неизворних говорника.” *Филолог–часопис за језик, књижевност и културу*, по. 21, 13–35. <https://doi.org/10.21618/fil2021013m>.
- Марковић, Јелена, and Ранка Станковић. у штампи. “Ја/ти/ми/ви у дискурсној компетенцији у светлу контрастивне анализе међујезика.” *Методички видици*.
- Ристовић, Зоран. 2012. “Од корпуса до учионице: примена паралелизованих текстова у настави енглеског језика у основној школи.” *ИНФОтека* 13 (2): 52–66.
- Ристовић, Зоран. 2016. “Кумулативни ефекти експлоатације вишејезичних корпуса у настави страних језик.” PhD diss., Универзитет у Београду, Филолошки Факултет.
- Шућур, Срђан Р. 2020. “Корпус као оруђе за проницање тајни међујезика.” *Радови Филозофског факултета: Часопис за хуманистичке и друштвене науке*, по. 22, <https://doi.org/10.7251/RFFP2022321S>.

Вики-библиотекар, пројекат за обуку библиотекара и студената за рад на Википедији

УДК 023-051: [030:004.738.5

САЖЕТАК: Вики-библиотекар је вишегодишњи пројекат обуке библиотекара и студената библиотекарства и информатике за коришћење вики алата укључујући писање чланака на Википедији. Пројекат постоји од почетка 2015. године, реализује га Универзитетска библиотека „Светозар Марковић“ уз логистичку и финансијску подршку Викимедије Србије. Сегмент пројекта који се бави обуком библиотекара званично је акредитован. Пројекат се проширио и на рад са студентима, као и рад са библиотекарицама изван самих обука кроз уређивачке маратоне, дигитализацију садржаја, учешће у другим вики активностима као што су 1Lib1Ref, Википедијанац стажиста. Активности пројекта су представљене на неколико конференција и парцијално у неколико публикација. Овим пројектом постиже се са једне стране боље упућивање библиотекара у информатичке могућности вики софтвера и начине како да своја знања што боље представе у дигиталном вики окружењу, док се са друге стране овим активностима значајно се повећавају текстуални ресурси Википедије на српском језику и њихов квалитет.

КЉУЧНЕ РЕЧИ: Википедија, обука библиотекара, Википедија у образовању, референце на Википедији

РАД ПРИМЉЕН: 30. мај 2021.

РАД ПРИХВАЋЕН: 19. јули 2021.

Ђорђе Стакић
djordje.stakic@ekof.bg.ac.rs
Универзитет у Београду
Економски факултет
Београд, Србија

Александра Поповић
popovic@unilib.rs
Оја Кринуловић
okrinulovic@unilib.rs
Универзитетска библиотека
„Светозар Марковић“
Београд, Србија

1. Увод

У ери када је интернет постао главни медиј који повезује све тачке света у реалном времену, онлајн енциклопедија Википедија¹ се јавља као незаобилазан информатички ресурс који тежи да на једном месту прикупи и учини доступним што већу количину знања насталог доприносом што већег броја људи. Идући у корак са тим трендом библиотекарски осавремењују свој рад и приступ изворима информација и корисницима. На тај начин као израз спремности да се нађе најбољи пут који обједињује Википедију и библиотекаре настао је пројекат Вики-библиотекар у организацији Универзитетске библиотеке „Светозар Марковић“ из Београда и Викимедије Србије.² Овај пројекат током више година свог рада изнедрио је велики број обучених библиотекара и студената библиотекарства за рад на Википедији, а резултовао је бројним написаним и допуњеним чланцима, постављеним референцама, креираним и дигитализованим фотографијама, скенираним старим књигама.

Већ од првих година постојања Википедије поставило се питање о односу библиотекара према Википедији. Према истраживању (Luyt et al. 2010) тај став је већином био позитиван. Према (Snyder 2013) библиотекарски имају већу свест о негативним него о позитивним странама Википедије, па је зато више користе у личне сврхе него за истраживање или помоћ корисницима. У истраживању (Vetter 2014) описано је како су студенти композиције на Универзитету у Охају користећи посебне збирке библиотеке креирали чланке за Википедију чиме се повећала мотивација студената и свест јавности о библиотечким ресурсима. У прегледном раду (Okoli et al. 2014) анализира се 99 научних радова који се у некој мери баве Википедијом и разматра се однос тих радова према читаоцима Википедије. У уобичајене кориснике Википедије ту се убрајају истраживачи, библиотекарски и студенти. У (Jemielniak and Aibar 2016) анализирају се перцепција и резерве академске заједнице према Википедији при чему се наводе могући сценарији мирног суживота. У раду (Faletar Tanacković, Đurđević, and Badurina 2015) приказано је истраживање на Филозофском факултету у Осиеку о искуствима и ставовима студената и наставника о употреби Википедије у универзитетском окружењу. Ту се показало да студенти

1. Википедија на енглеском језику, Википедија на српском језику, датум приступа 1. јул 2021.

2. Викимедија Србије, датум приступа 1. јул 2021.

имају позитивније мишљење о Википедији од наставника, као и да 19,8% наставника изричито забрањује коришћење Википедије студентима.

Занимљиву студију о цитираности чланака на Википедији у научним радовима и монографијама доноси (Kousha and Thelwall 2017). Узорак је имао 302328 научних радова и 18735 монографија индексираних у Скопусу³ у периоду од 2005. до 2012. године. Свега 5% научних радова међу својим референцама је имало чланак са Википедије, док је код монографија тај удео био 33%. У прегледним радовима тај удео је 8%. То говори о већој прихваћености Википедије као извора при писању књига него што је то случај код научних радова.

Према истраживању (Lewoniewski, Węcel, and Abramowicz 2017) на 7 већих Википедија (по броју чланака) и преко 10 милиона њихових чланака показало се да се значајан број референци понавља на више Википедија. Тако су највећи број заједничких референци имале енглеска и немачка Википедија, 345202, што је и разумљиво јер се ради о највећим енциклопедијама из тог скупа. Ако се у разматрање узму чланци који су присутни на бар 5 Википедија онда је највећи број заједничких референци на том скупу био између енглеске и руске Википедије. Према истраживању (Singh, West, and Colavizza 2021) на узорку од 29,3 милиона цитирања из 6,1 милиона чланака Википедије на енглеском језику установљено је да 6,7% чланака са Википедије цитира бар један рад из часописа са придруженим DOI,⁴ док је на Википедији цитирано око 2% укупних чланака који су индексирани у Web of Science⁵ и имају DOI. Истраживање (Pooladian and Borrego 2017) указује на недостатак стандардизације код навођења референци на Википедији где често недостају подаци или се наилази на грешке. У том истраживању установљено је да је свега 3% научних чланака из информатичких и библиотечких наука објављених у периоду 2001-2010 цитирано на Википедији до краја 2016. године. Узорак је обухватао 26542 научна рада из категорије „Information Science and Library Sciences“ индексираних у Web of Science.

3. Скопус (Scopus) је научна цитатна база која постоји од 2004. године, Scopus, датум приступа 1. јул 2021.

4. DOI је скраћеница од енглеског назива „Digital object identifier“ (идентификатор дигиталних објеката), користи се за јединствену идентификацију научних радова и других дигиталних ресурса; DOI, датум приступа 1. јул 2021.

5. Web of Science је водећа научна цитатна база; Web of Science, датум приступа 1. јул 2021.

2. Контекст настанка пројекта

Викимедија Србије основана је 3. децембра 2005. године као локални огранак Задужбине Викимедије.⁶ Њене активности су у великој мери окренуте сарадњи са образовним установама, у почетку сарадњи са факултетима и потом и са средњим школама. Већ у децембру 2005. представљена је у оквиру семинара Катедре за рачунарство и информатику Математичког факултета идеја да се и студенти могу укључити у писање чланака на Википедији. Оснивач и главни промотер ове идеје био је један од аутора овог текста, Ђорђе Стакић. До реализације те идеје је дошло већ наредне школске године када су се два факултета Универзитета у Београду, Факултет за физичку хемију и Филолошки факултет, укључила у образовне активности Викимедије Србије кроз писање чланака на Википедији уместо класичних семинарских радова. Први постављени чланак на Википедији на српском језику насловљен је ЕНИАС⁷ и посвећен је првом дигиталном електронском рачунару. Поставио га је студент Филолошког факултета 18. априла 2007. на предмету Интернет и веб технологије код професорке Цветане Крстев.

У даљем развоју образовног програма Викимедије Србије укључивали су се и други факултети и средње школе. На том развојном путу је априла 2012. формиран Академски одбор Викимедије Србије као група чланова која се фокусира на развијање и реализацију образовних активности повезаних са Википедијом. Детаљније о вики технологији, њеном развоју и значају, као и о пионирским подухватима развоја образовног програма на Википедији може се видети у раду (Stakić 2009). Резултати истраживања ставова студената о употреби Википедије као наставног средства на Педагошком факултету у Врању Универзитета у Нишу представљени су у раду (Stakić et al. 2021). Ово истраживање на узорку од 203 студента показало је да они позитивно оцењују примену Википедије у настави и да радије састављају чланке на Википедији него што пишу класичне семинарске радове.

Универзитетска библиотека „Светозар Марковић“ у Београду почела је са радом 24. маја 1926. године и данас је са око милион и 700 хиљада публикација једна од највећих научних библиотека у Југоисточној

6. Информације о огранцима, тематским организацијама и корисничким групама Викимедије: Wikimedia chapters, датум приступа 1. јул 2021.

7. ENIAC, датум приступа 1. јул 2021.

Европи. У оквиру библиотеке, почев од 2004. године, реализују се програми за обуку библиотекара који од 2014. године имају званичну акредитацију (Popović, Stolić, and Stanišić 2019). Универзитетска библиотека показује велику спремност и жељу да сарађује са другим организацијама које су посвећене заједничким циљевима. У оквиру таквих сарадњи одржана је трибина „Creative Commons - креативност и друштво знања“ у Универзитетској библиотеци 12. децембра 2013. године у заједничкој организацији Викимедије Србије и Универзитетске библиотеке.⁸ Током трибине су аутори чланка имали своја излагања: „Викимедија Србије и Универзитет у Београду у развоју слободног знања“ (Ђорђе Стакић) којим је представљен образовни програм и рад Академског одбора Викимедије Србије, као и „Creative Commons лиценце и е-тезе Универзитета у Београду“ (Александра Поповић и Драгана Столић). У оквиру трибине се родила идеја о томе да се успостави сарадња Универзитетске библиотеке и Викимедије Србије на подручју образовног програма и да се осмисли програм обуке библиотекара за рад на Википедији.

3. Развијање пројекта Вики-библиотекар

После првих планова и договора Универзитетске библиотеке и Викимедије Србије приступило се реализацији интерне радионице за обуку библиотекара Универзитетске библиотеке који ће представљати језгро учесника будућег пројекта и у највећој мери га касније реализовати. Ова обука одржана је 30. јануара 2014. у просторијама библиотеке. Током лета 2014. је написан предлог пројекта Вики-библиотекар Викимедији Србије која је потом одобрила средства за његову реализацију у оквиру свог годишњег плана за 2015. годину. У приближно исто време пројекат је конкурисао и код Народне библиотеке за акредитовани програм сталног стручног усавршавања библиотекара. Тада су као аутори пројекат поднели Александра Поповић, Оја Кринуловић и Ђорђе Стакић, а као установе Универзитетска библиотека „Светозар Марковић“ и Математички факултет Универзитета у Београду. Крајем исте године пројекат је званично акредитован за 2015. годину. Касније је пројекат акредитован још три пута, једном на годину дана (2016), затим на период од две године (2017-2018) и потом на период

8. Агенда трибине „Creative Commons - креативност и друштво знања“, датум приступа 30. мај 2021

од три године (2019-2021). Слично је и код Викимедије Србије пројекат одобран за сваку наредну годину почев од 2015. до 2021. године уз одговарајући годишњи буџет којим се покривају трошкови пројекта.

Пред почетак активности, ради боље припреме за почетак пројекта, у децембру 2014. године организован је први једномесечни програм Викимедије Србије „Википедијанац стажиста“ који је реализован у Универзитетској библиотеци. У оквиру програма је ангажована студенткиња Јована Милошевић која је држала радионице за библиотекарe, писала нове чланке, бавила се дигитализацијом и постављала фотографије под слободном лиценцом на Викимедијину оставу.⁹

Од 2015. године почеле су реализације акредитованих радионица за библиотекарe које су спроводили акредитовани предавачи, чланови и реализатори пројекта. Радионице су реализоване као дводневни догађаји од по три сата који према акредитацији укупно носе 6 бодова. Првог дана се реализује обука за рад на Википедији у облику предавања и практичне радионице. Другог дана се пишу и уносе чланци на Википедији на српском језику. Библиотекари се охрабрују да припреме материјал за чланак, па се онда на лицу места чланци уносе и прилагођавају вики синтакси.¹⁰ Важан део целог процеса писања чланака је навођење референци из релевантних извора што доприноси проверљивости чланака на Википедији, а што је библиотекарима само по себи веома блиско. По завршетку обука сви полазници добијају сертификате, а на Википедији остају резултати у облику написаних чланака. На овај начин сви учесници обуке имају одређену добробит и мотивацију да у пројекту учествују. Будући да су теме слободне, библиотекари су имали прилику да пишу и о својим библиотекама, знаменитостима из места из којих долазе и о свему другом што је релевантно за Википедију и прелази праг значаја.¹¹ Обогаћени садржај Википедије на српском језику доприноси већој видљивости библиотека у онлајн окружењу, као и уопште већој видљивости библиотека као

9. Викимедијина остава (енгл. Wikimedia Commons) је репозиторијум медијских датотека под слободним лиценцама; Wikimedia Commons, датум приступа 1. јул 2021.

10. Вики синтакса је начин којим се обележава текст у Медијавики софтверу који се користи на Википедији; правила форматирања, датум приступа 1. јул 2021.

11. Википедија има смернице о значају; Википедија:Значај, датум приступа 1. јул 2021.

занимања. И сами библиотекари на тај начин имају прилику да се међусобно боље упознају, повежу и добију прилику да раде на нечему од ширег значаја.

Осим акредитованих дводневних радионица одржаване су и додатне радионице за обуку заинтересованих библиотекара које су служиле да се неке ствари доуче, провежбају, као и да се напише још већи број чланака. Тако је већ у 2015. години било 7 додатних радионица.

Следећа активност у оквиру пројекта били су Уређивачки маратони. Они су почели као једнодневни догађаји од неколико сати уживо у оквиру којих се окупе заинтересовани учесници - библиотекари који пишу и постављају чланке на Википедију. Искуснији библиотекари помажу почетницима, и уопште учесници једни другима помажу да се снађу око нових и мање познатих ствари. Пошто се активности одвијају у библиотеци, неопходна литература је при руци што олакшава рад на креирању чланака. Током 2015. године су организована три уређивачка maratona под називима „Отворена наука“, „Јован Цвијић: 150 година од рођења“ и „Библиотекари заједно на Википедији“. У 2016. години су организована два тематска уређивачка maratona „Научни часописи у отвореном приступу“ у оквиру којих је било укупно 4 радионице. Тада су позвани библиотекари установа које издају часописе и шире представници научних часописа у отвореном приступу који се издају у Србији. С обзиром да је одзив био неочекивано велики, поновљен је уређивачки маратон и реализовано је више радионица. На тај начин је Википедија постала богатија за шездесетак чланака из ове области, што је значајно допринело да информације о овим часописима постану видљивије и доступније ширем кругу читалаца. Током 2017. године је поновљен уређивачки маратон „Научни часописи у отвореном приступу“ кроз две радионице. Активности су настављене и 2018. године када су била два уређивачка maratona о научним часописима и један под називом „Моја улица“. Будући да је обуку прошао велики број библиотекара из разних градова, овај маратон је имао одзив код многих претходних полазника акредитованих радионица. Из тог разлога он је поновљен и 2019. године два пута.

Почев од 2016. године у оквиру пројекта Вики-библиотекар у Универзитетској библиотеци одржава се и пракса за студенте библиотекарства и информатике Филолошког факултета Универзитета у Београду. Ова пракса је везана за предмет Дигитални текст 2. У оквиру праксе студенти похађају обуке, практичне радионице, пишу чланке и на крају добијају поене који им се рачунају приликом формирања оцене

из предмета Дигитални текст 2. Ово је вишеструко корисно и за саме студенте који, поред обуке повезане са Википедијом и писањем чланака, могу да осете и доживе амбијент једне велике и значајне библиотеке као што је Универзитетска и на тај начин изблиза се упознају са самим радом у библиотеци, са тиме како цео процес изгледа „изнутра“, као и са запосленим библиотекарима. Ове обуке су углавном вишедневне и резултују новим чланцима на Википедији.

Од 2018. године свим побројаним активностима придружило се и учешће у глобалној кампањи 1Lib1Ref.¹² Циљ ове кампање је да мотивише библиотекаре да дају свој професионални допринос Википедији кроз унос референци у чланке. Сама кампања почиње 15. јануара на рођендан Википедије и траје двадесетак дана. У оквиру ове кампање, поред библиотекара прикључују се и други корисници који желе да уносе референце.

Почетак пандемије у марту 2020. године одразио се и на активности везане за овај пројекат. Више није било могуће уживо одржавати радионице нити за библиотекаре нити за студенте, као што није било могуће уживо одржавати ни уређивачке маратоне. Међутим то није зауставило активности на пројекту који је по својој природи успео да се прилагоди, да значајан део активности реализује онлајн и да на тај начин постигне планиране метрике. Тако је 2020. студентска пракса остварена слањем презентација, брошура и видео материјала уз консултације имејлом, док је 2021. обука студената била онлајн преко Зума,¹³ где су такође накнадне консултације биле имејлом. Током 2020. одржана су три вишедневна онлајн уређивачка маратона: „Српска хумористичко-сатирична периодика 19. и 20. века“, „Моја улица“, „Штампа мога краја“, као и 2021. уређивачки маратон „Књижевност за децу“. Разноврсност ових тема допринела је добром одзиву библиотекара. Током 2020. пажња је посвећена дигитализацији публикација у оквиру пројекта, а током 2020. и 2021. године значајан део активности био је везан за глобалну кампању 1Lib1Ref.

12. Глобална кампања 1Lib1Ref (скраћено од „један библиотекар, једна референца“ на енглеском); 1Lib1Ref, датум приступа 1. јул 2021.

13. Зум (енгл. Zoom) је софтвер за видео конференције који је експанзију доживео током пандемије 2020. године; званична веб страна Zoom, датум приступа 1. јул 2021.

4. Постигнути резултати

У оквиру пројекта Вики-библиотекар одржана су три једномесечна програма стажирања, Википедијанац стажиста, и то у децембру 2014, јулу 2016. и марту 2019. године. Прво стажирање резултовало је са 18 нових чланака на Википедији, 102 постављене фотографије и 5 скенираних књига са укупно 4554 стране. Током овог стажирања одржано је и 8 интерних радионица у Универзитетској библиотеци за 10 запослених.¹⁴ Током другог стажирања дигитализовано је 8 књига (2585 страна) и постављено на Викимедијину оставу.¹⁵ Треће стажирање допринело је Википедији са 16 нових, 13 допуњених, 13 сређених чланака, 158 фотографија и 10 књига (2607 страна).¹⁶

Кроз акредитоване радионице Вики-библиотекара прошао је велики број библиотекара из различитих градова Србије. Током 2015. и 2016. године било је по 6 дводневних радионица кроз које је прошло око 70 библиотекара годишње са приближно толико написаних или допуњених чланака. Током 2017. године било је 7 радионица, затим 2018. године — три и током 2019. године четири реализације акредитованог семинара. У Београду је реализовано највише радионица које је похађало преко 100 библиотекара. Осим тога акредитоване радионице су одржане у још 13 градова: Ужице, Нови Сад, Бања Врујци, Чачак, Смедерево, Врање, Ваљево, Лесковац, Обреновац, Топола, Крагујевац, Јагодина и Крушевац. Одржавање радионица Вики-библиотекара по различитим градовима Србије је значајно допринело децентрализацији активности. Кроз сваку од ових радионица ван Београда прошло је од 8 до 19 библиотекара из 3-17 библиотека по радионици. У овим метрикама смо под библиотекарима назвали све библиотечке раднике који су прошли обуке, укључујући, поред библиотекара, и књижничаре. Током 2015-2018 одржане су укупно 22 акредитоване радионице Вики-библиотекара кроз које је прошло укупно 298 полазника (Popović, Stolić, and Stanišić 2019). Током 2019. године обуку је прошло још 42 полазника кроз 4 акредитоване радионице. Када се узму у обзир не само места у којима је семинар гостовао, него уопште одакле су били полазници, долази се до броја од 39 градова и општина у Србији. Полазници семинара су били из разних типова установа што је збирно приказано у Табели 1. Осим званичних радионица Вики-библиотекара било је и додатних радионица

14. Детаљи стажирања, датум приступа 1. јул 2021.

15. Детаљи другог стажирања, датум приступа 1. јул 2021.

16. Детаљи трећег стажирања, датум приступа 1. јул 2021.

за библиотекарe, и то 7 током 2015, 3 током 2016, и по 4 током 2017, 2018. и 2019. године. На свакој од радионица су, осим обука, писани нови и допуњавани постојећи чланци на Википедији. Од 2020. године, као што је већ поменуто у одељку Развијање пројекта Вики-библиотекар, није било радионица уживо због епидемије вируса корона.

Установе	Број
Јавне библиотеке	34
Архиви	2
Музеји	6
Школе	15
Друштва – удружења	11
Високошколске установе	77

Табела 1. Подаци о броју установа које су похађале курс Вики-библиотекар од 2015. до 2021.

Током реализације уређивачких маратона, и у оквиру њих су одржаване радионице и предавања о Википедији уколико је то полазницима било потребно. Било да су уређивачки маратони одржани уживо или само онлајн, њихови резултати су такође обухватили већи број нових и допуњених чланака. Међу најуспешнијим маратонима су били:

- „Библиотекари заједно на Википедији“ из 2015. са 44 нова чланка и 6 допуњених чланака;
- „Научни часописи у отвореном приступу“ који је током 2016. произвео укупно 58 нових чланака и један допуњен чланак, а 2017. године — 25 нових и три допуњена чланка;
- „Моја улица“ током којег је написано 37 нових чланака и допуњен један чланак 2018. године, односно 44 нова чланка 2019. године.

Посебно успешни су били вишедневни онлајн уређивачки маратони из 2020. и 2021. године чије су метрике у броју чланака:

- „Српска хумористичко-сатирична периодика 19. и 20. века“: 71 нови чланак, 5 допуњених чланака;
- „Моја улица“: 59 нових и 5 допуњених чланака;
- „Штампа мога краја“: 91 нови чланак, 6 допуњених чланака;
- „Књижевност за децу“: 76 нових и 12 допуњених чланака.

Неки од наведених maratona су имали значајан допринос у постављању слободних датотека на Викимедијину оставу. Ту су се посебно истакла три уређивачка maratona из 2020. године када је укупно постављено око 800 датотека.¹⁷

Током студентске праксе 2016–2021. године обуку Вики-библиотекара је укупно похађало 143 студента и као резултат је у просеку добијено између 20 и 30 чланака годишње, укључујући значајан број постављених слободних датотека, углавном фотографија. На пример, у оквиру студентске праксе 2021. године на Викимедијину оставу је постављено 36 датотека.

Од 2018. године, када се Вики-библиотекар укључио у кампању 1Lib1Ref, сваке године постижани су све бољи резултати у погледу уноса референци. Тако је 2018. постављено 786 референци, онда 2019. још 1672, затим 2020. још 4052, док је 2021. године унето чак 17112 референци. Према броју унетих референци у кампањи 1Lib1Ref, заједница Википедије на српском језику је на глобалном нивоу 2019. године заузела треће место, 2020. друго место, а 2021. године прво место. Рекордне (по броју унетих референци) 2021. године од 10 корисника који су унели највише референци на глобалном нивоу њих 5 било је са Википедије на српском језику.¹⁸

Укупно је кроз све активности пројекта Вики-библиотекар сваке године почев од 2015. креирано или значајно допуњено по неколико стотина чланака на Википедији на српском језику.

Пројекат Вики-библиотекар је представљен на више међународних и домаћих конференција, као и на више локалних скупова библиотекара и Википедијанаца организованих разним поводима. На овај начин је постигнута већа дисеминација резултата пројекта, објављено је више постера и чланака, а то је донело и неке нове сарадње. Међу објављеним чланцима издвајају се (Popović, Ševkušić, and Stakić 2015; Krinulović, Stijepović, and Stakić 2015; Antonić, Živanović, and Popović 2017; Popović et al. 2018). Постере са међународних конференција INFORUM – 22. годишње конференције о професионалним информационим ресурсима (енгл. 22nd Annual Conference on Professional Information Resources) (Праг 2016), INFORUM – 23. годишње конференције о професионалним информационим ресурсима (енгл. 23nd Annual Conference on Profes-

17. Активности и резултати Вики-библиотекара за 2020. годину, датум приступа 1. јул 2021.

18. Локални резултати кампање 1Lib1Ref за 2021. годину, датум приступа 1. јул 2021.

sional Information Resources) (Ппар 2017), 10. међународне конференције Европске смјернице за сарадњу библиотека, архива и музеја (Сапајево 2018) доносимо као слике у прилогу (слике 1, 2, 3).¹⁹ У раду (Obradović et al. 2020) Вики-библиотекар је поменут као један од пројеката Викимедије Србије, док се у делу о примени Creative Commons лиценци у Србији помиње већи број факултета, али и Универзитетска библиотека „Светозар Марковић“.

Значај пројекта Вики-библиотекар огледа се и у томе што су се њиме одшкринула врата и других библиотека које су желеле да учествују у активностима повезаним са Википедијом кроз ослобађање и дигитализацију садржаја, као и реализацију пројекта Википедијанац стажиста. После Универзитетске библиотеке најплоднија сарадња постигнута је са Библиотеком града Београда где је, осим пројеката библиотечког типа, Милица Буха развила пројекат Вики-сениор (почев од 2019. године) намењен обуци и активирању старијих полазника, пензионера, за рад на Википедији.

Још један значајан аспект Вики-библиотекара је помоћ у превазилажењу родног јаз на Википедији (Ratković and Madžarević 2021). Према разним истраживањима која су спровођена на узорку уредника Википедије, удео жена међу њима је у најбољем случају до једне шестине или једне петине (Graells-Garrido, Lalmas, and Menczer 2015; Maljković 2015). Пројекат Вики-библиотекар доприноси значајном повећању учешћа жена на Википедији, с обзиром да је и међу библиотекарима и међу студентима библиотекарства обрнут удео жена у односу на просек Википедије, већином са преко 80 процената. Иако ово није ставка од суштинског значаја за сам пројекат, на укупном плану ова чињеница има своју вредност и тежину. Сами чланови пројекта Вики-библиотекар од 2018. учествују у глобалној кампањи WikiGap²⁰ која се обично реализује у облику уређивачког маратона на Википедији и има за циљ да смањи родни јаз на Википедији.

Викимедија Србије је крајем 2017. године прогласила пројекат Вики-библиотекар за најбољи свој пројекат те године. Осим тога Викимедија Србије је доделила Александри Поповић годишњу награду „Бранислав Јовановић“ 2020. године како је наведено у образложењу награде, „За успешно вођен пројекат Вики-библиотекар и укупан рад и уређивање страница на Википедији на српском језику“.

19. Прва два постера доступна су и на линковима Постер Inforum 2016 и Постер Inforum 2017, датум приступа 1. јул 2021.

20. WikiGap, датум приступа 1. јул 2021.

5. Закључак

Дигитализација библиотечког фонда омогућава приступ различитим врстама садржаја. Пројекат Вики-библиотекар је и формиран из тог разлога и опстао је дуги низ година. Идеја водила пројекта је обogaћивање Википедије, као најважнијег пројекта Викимедијине фондације, са што више квалитетног садржаја. Библиотеке држе непроцењиво благо које није лако доступно свима. Отпремање садржаја на Викимедијину оставу идеалан је начин промоције дела са истеклим ауторским правима или оних чији аутори желе да им омогуће слободан приступ. Библиотекари — као добро повезана заједница чији чланови имају приступ референтној литератури на свом радном месту, вешти су у проналажењу валидних и проверљивих информација и упознати са богатим фондом и збиркама својих институција и имају искуство у обуци корисника различитих узраста и нивоа образовања тако да представљају савршен избор за уреднике Википедије. Ослањајући се на стручност, библиотекари могу значајно допринети бољем структурирању података и ефикаснијем индексирању садржаја у пројектима Википедије.

Пројекат Вики-библиотекар је за седам година постојања кроз разне облике активности обучио неколико стотина библиотекара из целе Србије за рад на Википедији. Осим тога кроз пројекат је прошло више генерација студената библиотекарства Филолошког факултета. У оквиру пројекта су реализовани бројни уређивачки маратони којима је значајно обogaћен садржај Википедије на српском језику. На тај начин постављено је или значајно допуњено неколико хиљада чланака. Осим тога већи број публикација је дигитализован и постављен на Викимедијину оставу, где се осим тога налази и већи број фотографија које су креирали и поставили учесници пројекта. Посебно успешна активност Вики-библиотекара била је учешће у глобалној кампањи 1Lib1Ref у којој је заједница Википедије на српском језику 2021. године унела највећи број референци на глобалном нивоу. Пројекат доприноси смањењу родног јаза на Википедији јер стицајем околности велика већина библиотекара који су прошли кроз пројекат су жене. Досадашње активности пројекта су делимично представљене у неколико публикација и постер презентација са конференција. У овом раду су те активности представљене потпуније и студиозније на једном месту.

Литература

- Antonić, Sanja, Igor Živanović, and Aleksandra Popović. 2017. “Altmetrics - Useful Way for Scientific Communication.” Rad predstavljen na IN-FORUM 2017: 23rd Annual Conference on Professional Information Resources, May 30–31, 2017, Prague. <https://www.inforum.cz/proceedings/2017/24/>.
- Faletar Tanacković, Sanjica, Anja Đurđević, and Boris Badurina. 2015. “Wikipedija u akademskom okruženju: stavovi i iskustva studenata i nastavnika.” *Libellarium: časopis za povijest pisane riječi, knjige i baštinskih ustanova* 8 (2): 161–199. <https://doi.org/10.15291/libellarium.v8i2.234>.
- Graells-Garrido, Eduardo, Mounia Lalmas, and Filippo Menczer. 2015. “First women, second sex: Gender bias in Wikipedia.” In *Proceedings of the 26th ACM conference on hypertext & social media*, 165–174. <https://doi.org/10.1145/2700171.2791036>.
- Jemielniak, Dariusz, and Eduard Aibar. 2016. “Bridging the gap between wikipedia and academia.” *Journal of the Association for Information Science and Technology* 67 (7): 1773–1776. <https://doi.org/10.1002/asi.23691>.
- Kousha, Kayvan, and Mike Thelwall. 2017. “Are Wikipedia citations important evidence of the impact of scholarly articles and books?” *Journal of the Association for Information Science and Technology* 68 (3): 762–779. <https://doi.org/10.1002/asi.23694>.
- Krinulović, Oja, Mile Stijepović, and Đorđe Stakić. 2015. “Bibliotečki doprinos “srpskoj Vikipediji”.” *Glas biblioteke* 21:21–28. <https://cacak-dis.rs/izdanja/casopisi/glas-biblioteke/>.
- Lewoniewski, Włodzimierz, Krzysztof Węcel, and Witold Abramowicz. 2017. “Analysis of references across Wikipedia languages.” In *International Conference on Information and Software Technologies*, 561–573. https://doi.org/10.1007/978-3-319-67642-5_47.
- Luyt, Brendan, Yasmin Ally, Nur Hakim Low, and Norah Binte Ismail. 2010. “Librarian perception of Wikipedia: Threats or opportunities for librarianship?” *Libri* 60:57–64. <https://doi.org/10.1515/libr.2010.005>.

- Maljković, Filip. 2015. *Izveštaj o rodnoj strukturi na Vikipediji na srpskom jeziku*. Wikimedia Commons, April 27th 2015. https://commons.wikimedia.org/wiki/File:Rodna_struktura_na_Vikipediji_na_srpskom_jeziku.pdf.
- Obradović, Ivan, Ranka Stanković, Marija Blagojević, and Danijela Milošević. 2020. "Open Educational Resources in Serbia." In *Current State of Open Educational Resources in the "Belt and Road" Countries*, 175–194. https://doi.org/10.1007/978-981-15-3040-1_10.
- Okoli, Chitu, Mohamad Mehdi, Mostafa Mesgari, Finn Årup Nielsen, and Arto Lanamäki. 2014. "Wikipedia in the eyes of its beholders: A systematic review of scholarly research on Wikipedia readers and readership." *Journal of the Association for Information Science and Technology* 65 (12): 2381–2403. <https://doi.org/10.1002/asi.23162>.
- Pooladian, Aida, and Ángel Borrego. 2017. "Methodological issues in measuring citations in Wikipedia: a case study in Library and Information Science." *Scientometrics* 113 (1): 455–464. <https://doi.org/10.1007/s11192-017-2474-z>.
- Popović, Aleksandra, Oja Krinulović, Mile Stijepović, and Đorđe Stakić. 2018. "GLAM projekat: galerije, biblioteke, arhivi i muzeji na internetu." In *Zbornik Radova – Asocijacija informacijskih stručnjaka, bibliotekara, arhivista i muzeologa (BAM)*, 10:191–193. <https://www.ceeol.com/search/article-detail?id=846203>.
- Popović, Aleksandra, Milica Ševkušić, and Đorđe Stakić. 2015. "Biblioteke i Vikipedija zajedno na webu: slobodno znanje za sve." *Digitalna humanistika: tematski zbornik u dve knjige, knj. 1*, 151–161. <https://doi.org/10.18485/dh.2015.1.ch12>.
- Popović, Aleksandra, Dragana Stolić, and Dejana Kavaja Stanišić. 2019. "Doprinos Univerzitetske biblioteke "Svetozar Marković" stalnom stručnom usavršavanju zaposlenih u bibliotečko-informacionoj delatnosti." *Čitalište* 18 (34): 32–39. <https://doi.org/10.19090/cit.2019.34.32-39>.
- Ratković, Nebojša, and Ivana Madžarević. 2021. "Women's Participation in Wikipedia: Cross-Border Balkan Perspective." *Área Abierta* 21 (2): 237–253. <https://doi.org/10.5209/arab.72763>.

- Singh, Harshdeep, Robert West, and Giovanni Colavizza. 2021. "Wikipedia citations: A comprehensive data set of citations with identifiers extracted from English Wikipedia." *Quantitative Science Studies* 2 (1): 1–19. https://doi.org/10.1162/qss_a_00105.
- Snyder, Johnny. 2013. "Wikipedia: Librarians' perspectives on its use as a reference source." *Reference and User Services Quarterly* 53 (2): 155–163. <https://doi.org/10.5860/rusq.53n2.155>.
- Stakić, Đorđe. 2009. "Wiki technology: Origin, development and importance." *Infoteka* 10 (1-2): 61a–69a. <http://infoteka.bg.ac.rs/index.php/en/archives/2009/infoteka-10-1-2-2009-69-78>.
- Stakić, Đorđe, Marija Tasić, Marko Stanković, and Milena Bogdanović. 2021. "Students' Attitudes Towards the Use of Wikipedia: A Teaching Tool and a Way to Modernize Teaching." *Área Abierta* 21 (2): 309–325. <https://doi.org/10.5209/arab.72760>.
- Vetter, Matthew A. 2014. "Archive 2.0: What composition students and academic libraries can gain from digital-collaborative pedagogies." *Composition Studies* 42 (1): 35–53. <https://www.jstor.org/stable/compstud.42.1.0035>.

Додатак

Open Science on Wikipedia: Libraries' Activities in Serbia

Aleksandra Popovic
popovic@unilib.bg.ac.rs

Sanja Antonic
Oja Krinulovic
Djordje Stakic



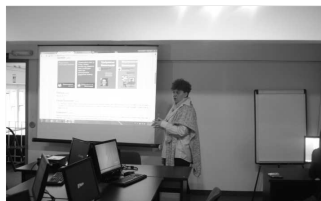
Open Science

Open science should make scientific information available. There is a tendency that scientific results are published immediately after the research. The application of open science has already shown that in that way science develops faster and raises the level of education of people. Guide to Open Science was published in 2014. Open science makes general science more transparent and accessible to a bigger number of researchers. Open science enables faster measurement of bibliometric indicators which influence both research and researchers (citation, h-index, Journal Impact Factor).

Wiki Marathon

Wiki Marathon is an event organized within Wiki Projects. Its topic is defined in advance and all participants either write or edit articles on that specific topic on Wikipedia. It can last a day or several days. Both trained Wikipedia editors and beginners can participate.

Wiki Marathon Open Science was organized within the project and articles on the topic were created on Wikipedia in Serbian. They are connected with the articles on Open Access Movement and Open Education, OA software, digital repositories and valuable sources of information. New articles were written also about the specific following topics: Budapest Open Access Initiative, Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, Cape Town Open Education Declaration and Bethesda Statement on Open Access Publishing.



Wiki Librarian

In 2015 the University Library "Svetozar Marković" in Belgrade, Serbia, started a very successful project Wiki Librarian in cooperation with the Faculty of Mathematics of the University of Belgrade and Wikimedia Serbia and the project has continued in 2016. The aim of the project is writing articles on Wikipedia in Serbian and editing pages on other Wiki Projects. Librarians are experts in finding valid and relevant sources of information. They have access to reference collections, commercial and OA e-journals, e-books from rich library collections and printed publications. Many libraries in Serbia joined the project whose aim is to put articles on various topics in OA and to make them available to everyone via Wikipedia.



Wikipedia

"Wikipedia is a free encyclopedia which everyone can edit". Wikipedia is used as one of the first sources of information. In order for an article on Wikipedia to be valid it has to be verifiable. That means that it must have appropriate references and additional literature. Wikipedia in Serbian was established in 2003. The title page was created in February 2003, and the first articles were published in September. Since then the number of articles has been increasing, and there is a trend in article extension and quality improvement. At the moment there are 291 languages on Wikipedia.

References

Guide to Open Science on link <http://digitheadslabnotebook.blogspot.rs/2014/01/guide-to-open-science.html>

Simcha Jong, Kremena Slavova (2014). When publications lead to products: The open science conundrum in new product development. Research Policy, Vol. 43, Iss. 4, Pages 645-654

image by Aleksandar Milosevic

22nd annual
CONFERENCE ON PROFESSIONAL INFORMATION RESOURCES

Слика 1. Постер са конференције INFORUM 2016

Periodicals on Wikipedia: Serbian Newspapers, Journals and Yearbooks

Aleksandra Popovic

popovic@unilob.rs

Sanja Antonic

Oja Krinulovic

Djordje Stakic



Wikipedia, Online Encyclopedia

Serbian periodicals are presented on Wikipedia which is a suitable platform for displaying valid and verifiable information available to users worldwide. Printing press reflects cultural level of a society and its role in the development of such a society is undisputed. Serbian periodicals have a long publication history as the first Serbian newspapers had appeared before the establishment and official proclamation of the restored Serbian state. The first Serbian journal in Slavonic-Serbian entitled "Slaveno-serbski magazin" was printed in Venice in 1768. The first scientific journal in medical sciences entitled "Srpski arhiv za celokupno lekarstvo" was published in 1872. For years bibliographies and lists of Serbian periodicals with tradition of over two centuries were being prepared. Today, this can be presented in a simpler and improved way via the biggest online encyclopedia which has been growing and developing on a daily basis.

Wikipedia, as a source of information, connects texts with factual data, illustrations, front pages, editors, authors, bibliographic data, different sources and literature thus making a unique whole and knowledge available to anyone. It also provides data via links to digital collections of newspapers and magazines.

periodicals
journals
newspapers

Wikipedia
Serbia
Online Encyclopedia



References

- Neil Selwyn, Stephen Gorard (2016). *Students' use of Wikipedia as an academic resource — Patterns of use and perceptions of usefulness*. The Internet and Higher Education, Vol. 28, Iss. 1, Pages 28-34.
- Ilija Reznik, Vladimir Shatalov (2016). *Hidden revolution of human priorities: An analysis of biographical data from Wikipedia*. Journal of Informetrics, Vol. 10, Iss. 1, Pages 124-131.

design by Aleksandar Milosevic



23rd annual
CONFERENCE ON PROFESSIONAL INFORMATION RESOURCES

Слика 2. Постер са конференције INFORUM 2017

Слика 3. Постеп са 10. међународне конференције Европске смјернице за сарадњу библиотека, архива и музеја, Сарајево 2018.

Друштвени медији и њихова улога у ширењу одређене представе о лепоти

УДК 316.62: 004.738.5

САЖЕТАК: Друштвени медији се из дана у дан користе све више и постају све популарнији. Друштвени медији су утицали на људе на много различитих начина, али најистакнутији су стандарди лепоте. Овај рад анализира ефекте који друштвени медији имају на начин на који људи сагледавају своје тело упоређујући га с оним што виде на друштвеним медијима. Анализира се, такође, како је овај утицај друштвених медија дорпинео ширењу нереалистичних апликација и филтера међу људима, посебно међу тинејџерима. Ове апликације и филтери су алати за уређивање слика који дозвољавају корисницима да промене своје слике тако да изгледају лепши. Ове преобликоване слике нису реалне и понекад могу водити слабом самопоштовању код младих жена. Сprovedена је мала анкета у којој је двадесет шест младих жена поделило своје ставове о искуству с друштвеним медијима. Циљ анкете је био да се сагледа како су друштвени медији променили њихове погледе на женско тело и стандарде лепоте.

КЉУЧНЕ РЕЧИ: друштвени медији, перцепција лепоте, апликације за улепшавање

РАД ПРИМЉЕН: 16. јун 2021.

РАД ПРИХВАЋЕН: 18. јул 2021.

Адиба Сидики (GL-7416)

zarasiddiqui86@gmail.com

Муслимански универзитет у

Алигару

Одељење за енглески

Алигар, Индија

1. Увод: коришћење апликација и филтера за улепшавање

Друштвени медији су свеprisутни у нашим свакодневним животима. Данас, ко год користи паметни телефон, склон је прикључивању некој

платформи друштвених медија. Према глобалном извештају Digital 2021 (Kemp 2021), раст друштвених медија се значајно убрзао од избијања пандемије Covid-19. Током прошле године број корисника је порастао за више од 13% с прикључивањем отприлике петсто милиона нових корисника, чиме је до почетка 2021. године укупан број корисника на глобалном нивоу достигао скоро 4,2 милијарде. Напредак дигиталних технологија одвија се убрзано. Друштвени медији имају данас огроман утицај на људе, посебно на младе. Једна од главних карактеристика Instagram-а и Snapchat-а је слање текстуалних порука као и дељење слика и видео записа. Популарност ових друштвених платформи је невероватно велика и свакодневно расте. Истраживање (Iqbal 2021b) показује да је највећи број корисника Instagram-а узраста 18–24, а да је Индија на другом месту по броју корисника.

Мода, нега коже и козметика, здравље и физичка спремност (Iqbal 2021a) су међу десет тема које највише интересују кориснике Instagram-а. Амерички модел Ким Кардашијан Вест је седма на листи особа које се највише прате на Instagram-у. Кендал Џенер, Клои Кардашијан и Кортни Кардашијан су међу двадесет највише праћених модела на Instagram-у. Све ове жене имају слично тело и карактеристике лице. Све оне имају истакнуту вилицу, високе јагодице, извајан нос и напућене усне. Чак и њихова тела су слична јер имају витке руке, ноге и струк са увећаном позадином и грудима. Ове жене су на вештачки начин постале лепе,¹ тако што су користећи пластичну хирургију потпуно измениле своја тела, примењујући увећање усана, Botox, операцију носа, уклањање нежељених длачица, и још пуно тога. Ако се погледају профили на Instagram-у модних инфлуенсерки у Индији попут Комал Пандеј, Дипа Булер Косла, Критика Курана, Ашна Шроф, Пуђа Мундра, Сантоши Шети, примећују се велике сличности у погледу заузетих поза и шминке. Њихове усне су напућене, нос извајан, вилица истакнута, јагодице изражајне а тело је крушколиког облика. Да би неко био леп, ове особине су постале обавезне. То је створило осећај униформности међу профилима на друштвеним медијима.

Лепота је субјективна и њена дефиниција се разликује од особе до особе. Може се видети да друштвени медији играју важну улогу у одлучивању какав тип тела се квалификује као леп. Стандардни тип лепоте се у нашем друштву дефинише кроз оглашавање, маркетинг и

1. Видети чланак Мехере Бонер “These Before-and-After Pics of the Kardashians Will Blow Your Mind” (Ове слике Кардашијанки пре-и-после ће вас распахметити) из *Costopolitan* (9.12.2020) Чланак

масовну репродукцију слика чиме се нормализује одређена представа о лепоти. У свему овоме постоји плурализам, али у њему су одређене особине незаобилазне. У данашње време су напућене усне, увећане очи, истакнута вилица, извајан нос, високе јагодице и затегнута кожа постали незаобилазни, посебно на друштвеним платформама. Дакле, постоје многе апликације које се баве овим суштинским предусловима лепоте, што се уочава и у филтерима Snapchat-а и Instagram-а. Људи који не могу да приуште скупу шминку и хируршке интервенције могу да мењају своје слике користећи ове апликације и филтере да би се уклопили у нормативе лепоте. Види се да их многе младе жене користе да би увећале усне и очи, промениле и тен и добиле затегнутије лице.

Модни часописи као што су *Vogue*, *Elle*, *Allure*, *Femina*, *Cosmopolitan*, итд., већ годинама имају битну улогу у дефинисању представе о томе шта је лепо. Сlike познатих личности и модела које се објављују у овим часописима обожавају многи и оне утврђују одређену представу о лепоти. Ове слике се често „фотошопирају“ или уређују да би жене изгледале савршено. Ова идеја о „фотошопирању“ је настала пре више од тридесет година (Lindblad 2020), када су 1987. године два брата, Томас и Џон Нол развили Photoshop и продали 1988. године лиценцу за његову дистрибуцију компанији Adobe Systems Incorporated. То је био први софтвер за уређивање слика понуђен широком кругу корисника. Photoshop се током година развио кроз многе нове верзије што је подигло уређивање слика на виши ниво. Постепено су пуштене у рад многе једноставније апликације за паметне телефоне, попут Beauty Plus, Pic-sArt, B612, Snapchat, YouCam, Beauty Booth, Photo Wonder, итд.

Технолошки напредак је био изузетан током последњих година. Апликације и филтери за улепшавање су у великој мери променили представу коју људи имају о лепоти. Људи их користе изузетно много да би се улепшали и да би на виртуалним платформама пројетовали своју симулирану слику. Француски филозоф Жан Бодријар је у свом делу „Симулакруми и симулација“ (Baudrillard 1994) анализирао везу између стварности, симбола и друштва. Његово је мишљење да је данашње друштво „заменило целокупну реалност и значење са симболима и знацима, а да је људско искуство симулација реалности.“ Бодријар је дефинисао симулацију као имитацију која замењује стварни свет. Он говори о одговорности савремених медија за замагљивање разлике између производа који су стварно потребни и оних за којима потребу ствара маркетиншка индустрија. За постмодерно друштво он каже „да више није у питању имитација већ замена стварног знаковима стварног.“

Ова идеја о перцепираној стварности остварена је коришћењем апликација и филтера за уређивање слика које су омогућиле људима да изгледају лепше него обично. Људи користе ове филтере и апликације да би створили хипер-реалну слику о себи на друштвеним платформама. Концепт хипер-реалног се дефинише као нешто што је „стваније од стварног.“ На пример, филмске звезде у филмовима изгледају гламурозно па се верује да су заиста такве. Али њихова стварна слика се разликује од оне коју приказује камера. На пример индијски глумац Рађникант, који игра највише у јужноиндијским филмовима, у стварности изгледа старије него као филмска звезда на биоскопском платну. Тако су људи створили хипер-реалну слику о себи да би се допали публици.

Ови филтери и апликације могу учинити чудо од слика људи. Може се додати шминка, посветлити кожа, сузити нос, додати лажне, веће трепавице, ставити руж за усне, увећати усне, груди и позадину. Популаран је тренд међу корисницима да постављају на друштвене платформе своје кристално чисте слике. Помоћу ових филтера и апликација можете лако да блистате на сликама иако сте можда тек устали из кревета. Да би ово илустровали, једна од учесница анкете је пристала да се њена слика користи у раду (табела 1). Разлике између ове две слике се јасно уочавају. Кожа је глаткија, нос ужи, усне напућене, груди повећане. За ове промене коришћена је апликација Beauty Plus.



Табела 1. Лево: фотографија пре измена; десно: фотографија после измена.

Занимљива ствар у вези ових филтера и апликација је да се у њима огледа еманципација више него у хируршким операцијама. Особа може да остане у својој кожи, а да коришћењем ових апликација измени лице и тело да би пројектовала своју симулирану слику виртуалним масама. Ова идеја је привлачна људима који постављају своје слике ради финансијске добити или просто да би их публика више ценила. Многи инфлуенсери на Instagram-у користе ове филтере и апликације да би створили лепше слике на виртуалним платформама што им понекад помаже да стекну огромну популарност – на пример, више следбеника и лајкова – а на крају, и више новца.

Прилично је фасцинантно да постоји толико људи који користе ове филтере и апликације да би се уклопили у одређену представу о лепоти не би ли стекли што већу популарност, а тиме и више новца. Тржиште инфлуенсера расте из дана у дан, а брендови компензирају више новца него икада. Многи инфлуенсери промовишу производе различитих брендова чиме зарађују огроман новац. Брендови обично придружују инфлуенсерима који промовишу њихове производе линкове или провизије засноване на промотивном коду. Постоје различити фактори од којих зависи зарада, попут квалитета произведеног садржаја, броја следбеника, стопе ангажованости профила (којом се мери колико следбеници профила комуницирају са објавом), итд. *Instagram Marketing Hub* је један од калкулатора (Schaffer 2021) којим се мере профили одређених инфлуенсера. Овај конкретан калкулатор се фокусира на три фактора: број следбеника, просечан број „лајкова“ по објави и стопа ангажованости. Може се, дакле, закључити да су филтери и апликације стекли огромну популарност међу корисницима Instagram-а, а инфлуенсери су на неки начин допринели да се њихово коришћење повећа.

Табела 2 приказује детаље профила неколико индијских инфлуенсера на Instagram-у који су са више детаља представљени у тексту који следи. Детаљи су преузети са веб странице *Instagram Marketing Hub* (Geyser 2020).

1. Индијска ТВ глумица, блогерка и друштвена инфлуенсерка на Instagram-у која прави кратке и духовите видео клипове (до 15 секунди) у којима користи неки од филтера за побољшање позадине и улепшавање лица. Она је уз помоћ операција добила перфектне напућене усне и узак нос. Има око деведесет седам хиљада пратилаца, а садржај који објављује тиче се углавном забаве и

р.бр.	број следбеника	просечно лајкова	ангажованост по објави	просечна зарада по објави у USD
1.	97,277	2,850	2.95%	\$390–\$650
2.	12,412	903	7.67%	\$84.75–\$141.25
3.	8,559	438	5.25%	\$79.5–\$132.5
4.	24,767	849	3.43%	\$153.75–\$256.25
5.	69,367	1,136	1.72%	\$277.5–\$462.5
6.	596,535	21,164	3.58%	\$1,782–\$2,970

Табела 2. Детаљи о профилима шест инфлуенсера на Instagram-у

хумора. Према њеном извештају са *Instagram Marketing Hub* зарађује између 29.098,10 и 48.496,83 индијских рупија (INR) по објави.

2. Физиотерапеуткиња и блогерка која себе назива Instagram инфлуенсерком. Користи филтере у готово сваком клипићу да би посветлила кожу. Промовише различите производе попут витминских додатака и производа за негу коже чувених међународних компанија. Има више од дванаест хиљада пратилаца и зарађује 6.323,4–10.538,73 INR по објави.
3. Модна инфлуенсерка која зарађује између седамдесет девет до сто тридесет два долара по објави (5.931,53 – 9.885,89 INR), а прати је више од осамдесет хиљада људи. Већина њених слика је филтрирана.
4. Четврти пример је ауторка videa на Instagram-у чији садржај се тиче козметике, а пре свега шминке. Има доста пратилаца а зарађује 11.471,36–19.118,94 INR по објави.
5. Једна инфлуенсерка на Instagram-у поприлично користи филтере у својим плесним клипићима. Она промовише различите производе за негу коже и косе, накит, одећу и сл. Прати је пуно људи а зарађује око 20.704,5–34.507,36 INR по објави.
6. Модел и инфлуенсерка са Instagram-а, која је уз то и TikTok краљица, себе назива Барбиком (Barbie Doll). Има огроман број обожаваалаца, а њен профил прати више од пола милиона људи. Зарађује по објави 132.955,91–221.593,18 INR што је јако пуно новца. Уз помоћ шминке и филтера она на клипићима и сликама личи на лутку Барбику. У њеним објавама се лако може уочити интензивно коришћење филтера. Понаша се и говори попут Барбике са замрзнутим пртама лица, налик на робота, као да је стварно у питању лутка Барби.

2. Прикупљање података и методологија

Истраживачки рад се заснива на примарном извору података прикупљених анкетом која је послата младим женама које користе платформе друштвених медија. Главни циљ упитника био је да се утврди колико ових жена користи апликације и филтере за улепшавање за своје слике. Такође требало је утврдити да ли коришћење ових филтера и апликација доводи до смањивања самопоштовања код жена које их користе и упоређују резултате са својом стварном кожом и телом. Упитник је послат насумично преко WhatsApp Messenger-a и Instagram-a на адресе тридесет жена, али није свима било пријатно да одговарају на ова питања. Двадесет шест жена је одговорило па је анализа обављена на основу тих података.

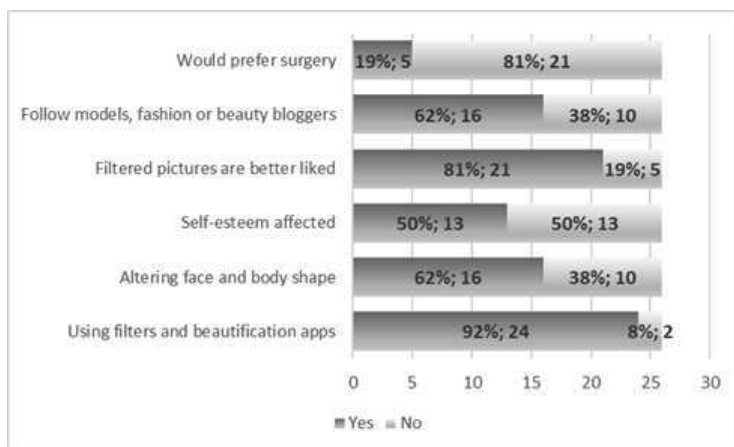
Садржај упитника је:

1. Како се зовете?
2. Колико имате година?
3. Чиме се бавите?
4. Каква је Ваша представа о лепоти?
5. Какав мобилни телефон користите?
6. Да ли пратите неке моделе на платформама друштвених медија? Ако је одговор потврдан, наведите које.
7. Да ли користите неке апликације или филтере за улепшавање за своје слике?
8. Да ли мислите да су Ваше филтриране слике допадљивије од нефилтрираних?
9. Да ли је коришћење апликација и филтера за улепшавање променило Вашу представу о лепоти?
10. Да ли је њихова примена утуцала на Ваше самопоштовање?
11. Шта обично мењате на свом лицу и телу када користите ове апликације или филтере?
12. Када бисте имали довољно новца, да ли бисте разматрали могућност операције ради промене неког дела лица или тела?

Највећи број одговора је стигао од студенткиња који спадају у старосну групу 18–24 из које је, како је речено на почетку овог рада, и највећи број корисника Instagram-a (Iqbal 2021a). Старосна група 18–24 означава у животу неке особе период преласка из адолесценције у зрело доба. Лако се уочава да су жене на преласку у зрело доба заинтересованије за свој изглед од старијих жена. Међу учесницама

је и неколико запослених жена и једна домаћица. Истраживање је намерно обухватило само жене јер је коришћење апликација и филтера за улепшавање популарније међу женама него међу мушкарцима, премда се у последње време може уочити да младе мушкарце такође привлачи њихово коришћење.

Графичка репрезентација података прикупљених према поменутим категоријама дата је на слици 1.



Слика 1. Анализа одговора на питања 6, 7, 8, 10, 11, 12

3. Анализа прикупљених података

На основу прикупљених података може се рећи да су се многе жене осећале несигурно због нереалистичних стандарда лепоте што их је навело да користе филтере и улепшавање. Двадесет четири од двадесет шест жена је рекло да користи филтере или неку врсту апликације за улепшавање да би посветле кожу или додале позадинске ефекте. Све ове жене су сагласне да се њихове слике боље примају када их филтрирају пре објаве. Ови стандарди лепоте су створили такве представе о лепоти да жене осећају да морају да промене своје тело да би их задовољили. Напућене усне су у тренду ових дана па је једна учесница анкете чак спремна на операцију да би повећала горњу усну. Савршена вилица и

високе јагодице се такође траже, па већина жена признаје да обично мењају вилицу коришћењем ових апликација и филтера.

Шеснаест учесница прати моделе и друге модне или козметичке блогерке које промовишу одређен појам лепоте који подразумева савршену вилицу, напућене усне, увећане очи и уста, високе јагодице и витак струк. Ови модели су углавном „пластични“, али су створили норматив лепоте који је постао идеал за већину жена. Млађе жене обожавају познате личности и моделе и диве им се због њихових савршених лица и тела. С тога већина младих жена осећа потребу да личи на њих, па прибегавају филтерима и апликацијама да би променили своје слике. Нос обликован у складу с модом се веома тражи, па би пет учесница спремних да се подвргну операцији најрадије одабрало риноплатику којом се реконструише нос и мења његов изглед.

Упитник је садржао питање које се односило на поимање лепоте испитаника. Већина њих је одговорила да је лепота нешто што особа носи у себи. Такође, самоуверена особа изгледа лепа многим ученицама анкете, док у исто време оне признају да користе апликације и филтере да би се улепшале, што изгледа контрадикторно. Такође, патријархално виђење жене као лепог објекта је покренуло употребу филтера и апликација. У патријархалном окружењу женино постојање се ограничава на потребе мушкараца, па је идеја да је поштена жена лепа жена натерала многе жене да се учине привлачним мушкарцима. Ово је нездрав начин размишљања који приморава млађе жене које је лако убедити да прибегну шминки, филтерима, апликацијама за улепшавање и у крајњем случају операцији како би изгледале лепо у складу са нормативима лепоте.

Многе учеснице су изјавиле да је њихово стално присуство на друштвеним медијима довело до тога да се упоређују са сликама лепотица, а такође су се сложиле да је коришћење Snapchat и Instagram филтера променило њихово мишљење о томе шта је лепо. Једна од учесница је рекла да ови филтери које користи толико много људи омаловажавају оне којима ионако недостаје самопоуздање и непријатно им је да покажу стварни изглед своје коже. То ствара код посматрача илузију о изузетно светлој и чистој кожи, уском лицу и носу и још којешта. Овакве слике слушају људе јер они почињу да упоређују филтриране слике са својом кожом у стварности. Филтери и апликације за улепшавање су потпуно променили поимање лепоте јер помоћу њих сви недостаци нестају. Може се учинити да све нестане. Кожа може постати глаткија па младе жене често мисле да би и њихова кожа

требало да буде беспрекорна и да уопште не би требало да имају поре. А то очигледно ствара токсично октужење и токсичне ствари које жене приморавају да поверују да је особа коју гледате кроз филтер боља од особе коју видите у огледалу.

Самопоштовање је још један важан део у животу људи. Ради се о субјективној процени сопствене вредности. Истраживање је покушало да установи да ли су жене осетиле да поређење са филтрираним хипер-реалним сликама повређује њихово самопоштовање. Тринаест учесница је одговорило да је на неки начин њихово самопоштовање било повређено, посебно када су почеле да упоређују своје филтриране слике са својим изгледом у огледалу. Особа на филтрираној слици није стварна, али када су жене почеле да теже да баш тако изгледају то је потиснуло њихово самопоштовање. Многе су се сложили да је интензивно коришћење ових филтера учинило да им је сада немогуће да објаве нефилтрирану слику. Више се не осећају сигурне да објаве слике које нису дотеране. Неке су чак рекле да користе само фотоапарат с филтером уместо стандардног фотоапарата.

Ако се удубимо у психоаналитичку теорију, ова жеља за изгледом са филтриране слике и прибегавање операцији одражава се на *ид* особе, његов импулсивни и несвесни део. Према Фројдовој психоаналитичкој теорији (Freud 1990), структура личности је подељена у три дела: *ид*, која је првобитна и нагонска компонента чији је циљ задовољавање основних жеља и потреба; *супер-его*, носилац моралних норми; *его*, који се управља принципом реалности у посредовању између *ида* и *суперега*. Ове апликације и филтери омогућавају људима да створе нереалну слику о себи, а њихова жеља да баш тако изгледају и уклопе се у одређену представу о лепоти само да би задовољили свој *ид* види се у ефектима интензивног коришћења ових филтера и апликација. *Ид* ради на принципу задовољства и жели да све што се пожели буде и задовољено. Операција је радикална опција која је болна и може имати озбиљне последице. Ако и поред свега неко и даље тежи да операцијом постане налик на своју филтрирану слику треба да зна да је то проблематично и да се упушта у нешто опасно и несигурно. Ове операције су често неуспешне и њихов исход је ужасан. Дозволити да ваш несвесни ум превлада вашу рационалну личност терајући вас на операцију није уопште добра идеја.

4. Закључак

Утицај западне културе на Индију довео је последњих година до великих промена у стандардима лепоте. Идеја „што светлија то лепша“ је дубоко укорееена у индијско друштво. Брачни огласи су увек наглашавали светлији тон девојчина тена. Постоји чак и посебна потражња за девојкама светле пути које се помињу у огласима (Gelles 2020). Од свих девојачких квалитета, по стереотипима индијског друштва светли тон коже је неопходан да би се за девојку рекло да је лепа, а он се највише и тражи. „Колоризам“ је дубоко укорееен у индијски начин размишљања, посебно када су девојке у питању, што га чини родно пристрасним. Приход индијског тржишта крема за посветљавање коже наводно је износио 30 милијарди долара у 2019. години, а очекује се да ће прећи 50 милијарди долара до 2023. године.² Идеја да је светао тен суштински квалитет лепих жена урезана је у свест Индијаца. После тога, идеал лепоте у савременој Индији подразумева велике очи, црну дугу косу, пуне црвене усне, витак струк са ширим боковима (Majidi 2020). Ови неопходни услови се донекле разликују од нових глобалних трендова у поимању лепоте, јер, на пример, светла пут није више толико битна захваљујући расној једнакости. Такође, дуга и густа, валовита црна коса је природена индијском идеалу лепоте, док глобална тежња за напућеним уснама то није. Дискриминација и расизам нису присутни само међу заједницама већ и унутар сваке од њих. Предност која се даје светлијем тену присутна је у многим земљама Азије, Африке и Јужне Америке.³ Према истраживању (Sims and Hirudayaraј 2016), жене са тамнијим тенем имају у Индији и, донекле, у суседним земљама мање шансе да постану стјуардесе, новинарке, рецепционерке, манекенке, продавачице, глумице, и уште да се баве послом који захтева комуникацију с људима јер би биле оцењене као „непривлачне“. Чак шест до седам учесница анкете је директно или индиректно признало да свој тен или тон коже осветљавају помоћу

2. Видети чланак Pia Krishnankutty “Before Fair & Lovely, there was Afghan Snow – all about the fairness creams market in India.” (Пре светла и љупка постојао је авганистански снег – све о тржишту крема за посветљавање у Индији) у *Print* (26. јуни 2020) Чланак

3. Видети чланак Meera Estrada “Commentary: Shadeism is the Dark Side of Discrimination We Ignore.” (Коментар: сенчење (разликовање по нијансама) је тамна страна дискриминације коју игнорисемо) у *Global News* (објављено 24. маја 2019), Чланак

филтера или апликација за улепшавање. Већина анкетираних жена су или студенткиње или су запослене, па се, према томе, не може рећи да ниво или врста образовања имају везе се коришћењем филтера и апликација за улепшавање на друштвеним медијима. Такође, може се приметити да високо квалификовани људи, попут лекара, током текуће пандемије Covid-19 постављају плесне или музичке клипиће на платформама друштвених медија на које су обично применили филтере.

Може се, према томе, закључити да друштвени медији играју важну улогу у дефинисању одређених норми лепоте, јер сви врхунски модели и модни блогери изгледају слично, облаче се слично, па чак објављују сличне слике. Ови трендови наводе друге жене да их следе, при чему оне морају да потраже помоћ дигиталних апликација и филтера да би их имитирале. Данас створена слика лепог лица и тела дефинише битне услове без којих се не може бити леп. Ради се углавном о увећаним уснама и очима, глаткој кожи, високим јагодицама, уочљивој вилици, витком струку, истакнутој позадини, уском носу, итд. На многе жене је то утицало захваљујући значајној присутности друштвених медија данас. Модели и модни блогери су дефинисали данашње стандарде лепоте који приморавају друге младе жене да их имитирају. Осим тога, многи инфлуенсери користе филтере и апликације за улепшавање да би привукли пажњу и самим тим зарадили више новца. Дакле, многе жене користе филтере и апликације које им помажу да истакну вилицу, напуће усне, загладе кожу, учине струк виткијим, итд, што је неизоставно код неких жена смањило самопоштовање јер нису могле да се осећају задовољно са својом стварном кожом и телом. Такође, покушај да изгледају као на својим филтрираним сликама је код њих створило неку илузију, што све води осећају незадовољства сопственим телом и, у многим случајевима, смањивању општег самопоуздања. Негативан ефекат је да многе жене више не могу да објављују нефилтриране слике. А оне које у толикој мери теже да изгледају попут своје филтриране верзије да би се одлучиле на операцију показују какав је ментални утицај ових филтера и апликација.

Ово истраживање је било усмерено углавном на млађе жене. Оно би се могло проширити укључивањем младих мушкараца да би се утврдило да ли су и они под истим притиском да побољшају своје виртуалне слике на платформама друштвених медија, а такође и на жене старијих генерација. Стални притисак који друштвени медији врше на кориснике да се уклопе у њихов норматив лепоте утицао је на поимање лепоте код људи. Такође, овај снажан порив може нагнати људе да се одлуче на

операцију, што се одражава на њихов *ид*, то јест нагонски и несвесни ум. Проширивањем психолошког опсега истраживања овај аспект – надвладавање *ида* над рационалним ја – могао би се дубље истражити.

Захвалност

Овај рад је настао у току ауторкиних мастер студија енглеске књижевности на Муслиманском универзитету у Алигару у Индији.

Литература

- Baudrillard, Jean. 1994. *Simulacra and simulation*. University of Michigan press.
- Freud, Sigmund. 1990. *The Ego and the Id*. Edited by James Strachey. W. W. Norton & Company.
- Gelles, Rebecca. 2020. "Fair and Lovely: Standards of Beauty, Globalization, and the Modern Indian Woman, Independent Study Project (ISP) Collection." Accessed June 22, 2020. <http://thelionandthehunter.org/norms-of-beauty-in-india-fair-is-beautiful-a-legacy-of-colonialism-and-globalization/>.
- Geyser, Werner. 2020. "Influencer Marketing, in Instagram Marketing Hub." Accessed October 29, 2020. <https://influencermarketinghub.com/instagram-money-calculator/>.
- Iqbal, Mansoor. 2021a. "Instagram Revenue and Usage Statistics, Business of App." Accessed May 24, 2021. <https://www.businessofapps.com/data/instagram-statistics/>.
- Iqbal, Mansoor. 2021b. "Snapchat Revenue and Usage Statistics, Business of App." Accessed May 6, 2021. <https://www.businessofapps.com/data/snapchat-statistics/>.
- Kemp, Simon. 2021. "Digital 2021: Global Overview Report." Accessed January 27, 2021. <https://datareportal.com/reports/digital-2021-global-overview-report#:~:text=Social%5C%20media%5C%20user%5C%20%20numbers%5C%20increased%20by%5C%20the%5C%20state%5C%20of%5C%202021.>

- Lindblad, Mason. 2020. “The History of Photoshop – Photoshop Through the Years, Filtergrade.” Accessed September 22, 2020. <https://filtergrade.com/history-of-photoshop-through-the-years/>.
- Majidi, Khesraw. 2020. “Norms of Beauty in India Fair is Beautiful: A legacy of Colonialism and Globalization, The Lion and the Hunter.” https://digitalcollections.sit.edu/isp_collection/1145/.
- Schaffer, Neal. 2021. “How Much Should I Pay an Influencer? Testing Out 9 Instagram Influencers Calculators, in Influencer Marketing.” Accessed June 6, 2021. <https://nealschaffer.com/instagram-influencer-calculators/>.
- Sims, Cynthia, and Malar Hirudayaraj. 2016. “The impact of colorism on the career aspirations and career opportunities of women in India.” *Advances in Developing Human Resources* 18 (1): 38–53.

Инфотека (Q25460443) у Википодацима

УДК 004.62: [030:004.738.5

САЖЕТАК: *Википодаци* (енгл. Wikidata) су база знања Задужбине Викимедија која представља заједнички извор различитих врста података које користе не само други Википедијини пројекти, већ све више и бројне апликације семантичког веба. У овом раду ћемо презентовати пример интеграције Википодатака са дигиталним библиотекама и екстерним системима, као и могућност убрзања припреме и уноса података на примеру радова из часописа за дигиталну хуманистику *Инфотека*.

КЉУЧНЕ РЕЧИ: семантички веб, отворени повезани подаци, викподаци, Инфотека, метаподаци часописа

РАД ПРИМЉЕН: 24. јун 2021.

РАД ПРИХВАЋЕН: 16. јули 2021.

Ранка Станковић

ranka.stankovic@rgf.bg.ac.rs

Универзитет у Београду

Рударско-геолошки факултет

Београд, Србија

Лазар Давидовић

lazarmdavidovic@gmail.com

Универзитет у Београду

Београд, Србија

1. Увод

Википодаци (енгл. Wikidata)¹ су база знања Задужбине Викимедија која представља заједнички извор различитих врста података, како конкретних тако и апстрактних. Похрањене податке могу да користе и други Викимедијини пројекти, као што је *Википедија*, али и шира заједница за различите потребе, чиме доприносе померању границе од машинске читљивости ка машинској разумљивости података на вебу. У овом раду ћемо презентовати пример интеграције Википодатака са дигиталним библиотекама и екстерним системима, као и могућност убрзања припреме и уноса података на примеру радова из часописа за дигиталну хуманистику *Инфотека*.

Семантички веб је проширење постојећег веба где се информацијама даје прецизно дефинисано значење и који омогућава бољу сарадњу рачунара и корисника. Отвореност и делимична структурираност ресурса који се развијају у организацији Викимедије је била основа за

1. Викидата

изградњу бројних машински читљивих ресурса, какав је на пример *DBPedia*,² који се ослањају на стандардизоване језике семантичког веба. Концепт семантичког веба и технологије отворених повезаних података, проширују традиционални веб употребом стандардног језика за означавање и сродних алата за обраду, где управо RDF (Resource Description Framework), оквир за описивање ресурса на вебу, има велику улогу и омогућава ефикаснија решења за проналажење информација (Shah et al. 2002). Да би семантички веб функционисао, рачунари треба да имају приступ структурисаним колекцијама информација и да утврде дефинисана правила аутоматизованог управљања. Википодаци се управо уклапају у те трендове развоја информационих технологија, које померају границе од машинске читљивости ка машинској разумљивости података на вебу.

Пројекат *Scholia*³ (Nielsen, Mietchen, and Willighagen 2017) један од првих свеобухватних подухвата осмишљених да се библиографске информације, научни профили аутора и институције представе користећи Википодатке. Управо резултати овог пројекта и доступност радова *Инфотеке* у различитим дигиталним облицима су били инспирација за „викификацију“ радова часописа *Инфотека*. Видевши конкретно веб стране *scholiaEvent*⁴ и *scholiaTopic*,⁵ покренута је слична акција, да се на основу постојећих метаподатака о радовима из часописа

2 warnings *Инфотека* креирају повезани (RDF) подаци о ауторима и радовима, а на сајту дигиталне библиотеке *Библиша*⁶ (Stanković et al. 2015) уз радове из часописа додају линкови ка Википодацима и прикаже граф коауторства, као интерактивна страница. Имплементација представља студију случаја која се може даље проширити на друге часописе, као и на друге случајеве употребе, на пример, конференције и дигиталне библиотеке. Алат *Scholia* се развија у оквиру веће иницијативе *WikiCite*,⁷ која настоји да индексира библиографске метаподатке у Википодацима о ресурсима који се

2. DBPedia

3. Scholia, Scholia у Википодацима

4. Преглед протеклих и предстојећих конференција са подсетником са роковима за слање радова ScholiaEvent

5. Преглед научних и стручних радова, као и њихови аутори и теме које се заједнички јављају, груписани по тематским целинама: Википедија, машинско учење, биологија, храна и сл. scholiaTopic

6. Библиша

7. WikiCite

могу користити за поткрепљивање тврдњи изнетих у Википодацима, Википедији или негде другде. У времену када смо преплављени нетачним информацијама на вебу, одговарајуће поткрепљивање информација релевантним изворима сигурно игра важну улогу. Како смо желели да аутоматизујемо, колико год је могуће, процес припреме и уноса података, истражили смо различита решења од којих смо користили два: *OpenRefine*⁸ и *QuickStatements*,⁹ о чему ће више речи бити у наредним одељцима.

Сарадња *Викимедије Србије*¹⁰ са Универзитетом у Београду има дугу традицију (Stakić 2009). Универзитетска библиотека „Светозар Марковић“ заједно са Математичким факултетом Универзитета у Београду и Викимедијом Србије је 2015. године почела пројекат *Вики-библиотекар*, са идејом да се постави што више квалитетних садржаја на Википедију (Porović, Ševkušić, and Stakić 2015). Википодатке, као мрежу отворених података је користила Андоновски (2020) за описивање језичких ресурса, конкретно романа, који су саставни део српско-немачког литерарног корпуса (Andonovski, Šandrih, and Kitanović 2019). На Рударско-геолошком факултету се студенти више година обучавају за унос и коришћење Википодатака,¹¹ а на докторским студијама *Интелигентни системи*, при Универзитету се у оквиру предмета *Представљање знања* и *Семантички веб* истражују апликативне могућности отворених података. У оквиру COST акције CA16204 (2017-2021) „Удаљено читање за европску историју књижевности“¹² се ради на уносу метаподатака о српским романима из корпуса *srpELTeC*¹³ (Krstev et al. 2019) и повезивању Википодатака са различитим апликацијама, од којих је једна Аурора.¹⁴ Резултатима који се представљају овим

8. OpenRefine (претходно Google Refine) је алат за рад са неуређеним подацима: чишћење, претварање из једног формата у други, уз допуњавање екстерним подацима путем веб сервиса OpenRefine

9. Алат за уређивање ставки Википодатака: додавање и уклањање изјава, ознака, описа и сл. QuickStatements

10. Wikimedija

11. Унос и коришћење Википодатака

12. Један од најважнијих циљева ове акције је припрема вишејезичног корпуса (названог European Literary Text Collection - ELTeC) који ће, када буде потпуно завршен, садржати по 100 романа први пут објављених у периоду 1840-1920. за више европских језика.

13. srpELTeC

14. Аурора

радом су допринели и чланови Друштва за језичке ресурсе и технологије *JePTex*.¹⁵

2. Википодаци

Википодаци (енгл. Wikidata) су база знања чија је сврха да буде заједнички извор одређених врста података (нпр. број становника државе, место рођења, датум оснивања), које користе други Викимедијини пројекти као што је Википедија. У том смислу је слична Викимедијиној остави где се складиште медијске датотеке којима се приступа из других Викимедијиних пројеката. Википодаци су оријентисани на документе, усредсређени на ставке, које представљају теме, концепте или објекте. Свакој ставци додељен је јединствени, трајни идентификатор, позитиван цео број са префиксом великог слова Q, познатог као „QID“. Ово омогућава превођење основних информација потребних за препознавање теме коју ставка обухвата, а да се не фаворизује било који језик, са циљем да се обезбеди јединственост значења конкретнoг појма.

Наведимо неке примере ставки: места (Нови Сад: Q55630, Лондон: Q84, Звездара (Београд): Q12645852), особе (Ђорђе Балашевић: Q342045, Тим Бернерс-Ли: Q80, Хеди Ламар: Q49034), догађаје (Први српски устанак: Q368689, концерт: Q182832, маратон: Q40244), предмете (столица: Q15026, чаша: Q81727, тигањ: Q127666)), појмове (радост: Q935526, страх: Q44619, појам: Q151885), књижевна дела (*Горски вијенац*: Q1192476, *Дон Кихот*: Q480, *Игра престола* (књига): Q1751870), филмови (*Лепота порока*: Q4239792, *Коса* (филм): Q757156), серије (*Игра престола*: Q23572, *Ало, ало!*: Q425628), балет (*Дон Кихот* (балет): Q1239463)... Концепти који стоје иза ставки треба да буду јединствени, али се дешава да постоје две ставке под истим називом, Никола Тесла (Q9036) представља славног научника, а *Никола Тесла* (Q2732597) представља насеље (Q486972) у Нишкој Бањи (Q954986) које је по њему добило име. Препорука је да се код вишезначних ентитета у загради дај додатно појашњење, као што је претходно наведени *Дон Кихот* (балет) или *Игра престола* (књига). Дакле, ставка је повезана са јединственим идентификатором (QID), идентификатор је повезан са паром: насловом и описом, како би се уклонила било каква двосмисленост.

15. JePTex

Универзитет у Београду је државни универзитет и члан Европске асоцијације универзитета

Универзитет у Београду је основао Доситеј Обрадовић 1808.

Q240631 P31 Q875538.	Q240631 P31 Q875538; P463 Q868940; P112 Q347659; P571 „1808“.
Q240631 P463 Q868940.	
Q240631 P112 Q347659.	
Q240631 P571 „1808“.	
Доситеј Обрадовић је рођен 17. фебруара 1742 у Чакову, а умро је 7. априла 1811. Био је лингвиста, песник, писац и филозоф.	
Q347659 P569 „17. фебруар 1742“.	Q347659 P569 „17. фебруар 1742“; P19 Q325736; P570 „7. април 1811“; P106 Q14467526; P106 Q49757; P106 Q36180; P106 Q36180.
Q347659 P19 Q325736.	
Q347659 P570 „7. април 1811“.	
Q347659 P106 Q14467526.	
Q347659 P106 Q49757.	
Q347659 P106 Q36180.	
Q347659 P106 Q4964182.	

Табела 1. Примери ставки Википодатака

Идентификатор ставке (QID), осим што је повезан са насловом и описом, може имати више псеудонима и одређени број изјава (тврђења, израза) којима се представљају њена својства и вредности. Изјава је уређена тројка: (ставка, својство, вредност), где је ставка (Q) – било која тема (особа, предмет, место, концепт), својство (P). Релација¹⁶ или карактеристика релевантна за ставку може бити на пример: боја косе (P1884) за људе, издавач (P123) за публикована дела, оснивање (P571) за организације и сл. Вредност ставке може бити сам „литерал“ односно ниска карактера (на пример: дужина Дунава је 2860 км) или референца на неку другу ставку (на пример главни град Србије је Београд). Ставка може бити описана низом изјава од којих свака даје једну чињеницу или податак о ставки. У табели 1 је дато неколико примера реченица на природном језику и записивање истих информација на Википедији, у

16. У математици ако је уређени пар (x, y) у релацији ρ тада кажемо да је елемент x у релацији са елементом y и пишемо као тројку: $x\rho y$. Слично, ставке у Википодацима релације изражавају као тројке, тако да релација Тесла-начин живота-вегетаријанство се бележи као Q9036 P1576 Q83364.

виду тројки субјекат, предикат и објекат (лева колона) и у скраћеној нотацији (десна колона).

У претходној табели, иза примера на српском језику су у првој колони дате тројке, односно реченице облика субјекат-предикат-објекат. Ако желимо да се прецизније изразимо, рећи ћемо да су то RDF тројке, где је RDF скраћеница од Resource Description Framework (Q54872), односно оквир за описивање ресурса на вебу. Реченице се завршавају тачком. У другој колони је приказана скраћена нотација, којом се избегава понављање субјекта, тако да знак „“ упућује да се следећи предикат односи на исти субјекат.

Важна карактеристика Википодатака је да имају два лица: једно намењено људима и друго намењено машинама, што омогућава бројне примене у обради природних језика. Поменимо неке: класификација текста, индексирање, анализа текста, генерисање текста, сумаризација, нормализација, повезивање и сл. Друга важна карактеристика је вишејезичност, која омогућава да се свака ставка може повезати са називом на било ком језику који је регистрован у Викимедијиним ресурсима, што отвара пут бројним применама, почев од аутоматског превођења и класификације вишејезичних докумената, до анализе садржаја на вебу и друштвеним мрежама.

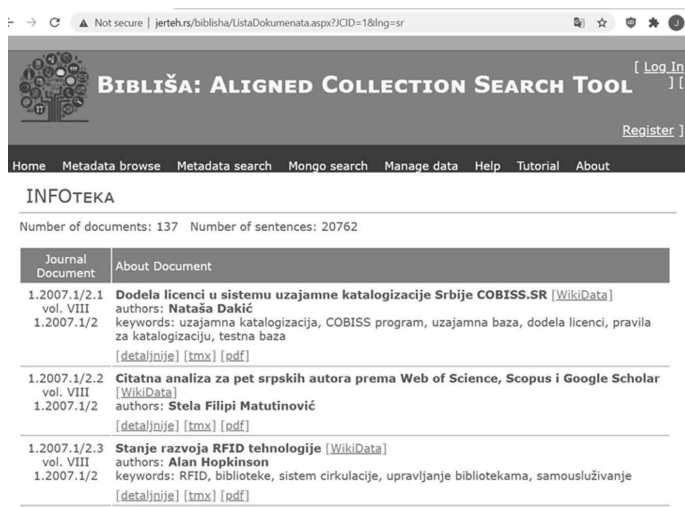
3. Аутоматизација уноса у Википодатке

Појединачни унос података је често временски захтеван посао, али се може убрзати када већ постоје подаци похрањени у различитим дигиталним форматима. Уз одговарајућу припрему, они се могу полуаутоматски увести у Википодатке. Дакле, основна идеја је да се убрза унос података, о резултатима истраживања у области дигиталне хуманистике у Србији и старих српских романа, како би се повећала видљивост уједно и српског језика, нашег културног наслеђа и резултата истраживања у Србији, а свакако отворио пут и за многе друге скупове података.

Да би аутоматизација уноса била могућа, први корак који је требало урадити јесте прикупљање и припрема података. Други корак се односи на избор назива ставки у подацима које ће се користити за идентификацију предиката и креирање шеме уноса. Шемом дефинишемо повезивање вредности са ставком, односно, субјектом помоћу предиката као посредника.

Иако је крајњи циљ унос података о радовима, унос њихових аутора је био неизоставан корак и предуслов даљем раду. По завршеном уносу креирани су SPARQL¹⁷ упити за различите приказе, користећи и интегрисане технологије у Википодацима за визуализацију резултата.

Сваки рад часописа Инфотека из двојезичне дигиталне библиотеке Библиша је повезан са припадајућим Википодацима, тако да може директно да се дође до конкретног, појединачног приказа у Википодацима, али и да се интегришу неки од корисних визуалних приказа у оквиру саме апликације. Слика 1 приказује пример панела у Библиши за приказана прва три рада у колекцији, где се види да се иза наслова сваког рада налази линк ка путањи ресурса у Википодацима.



Слика 1. Панел дигиталне библиотеке Библиша са прегледом метаподатака о Инфотеци

Први рад са листе би се превео на језик Википодатака на следећи начин:

17. SPARQL је протокол и упитни језик за RDF базе података који омогућава извлачење вредности, истраживање структуре података, спајања података који долазе из различитих база, трансформацију RDF података из једног речника или графа у други. Спецификација упитног језика SPARQL.

„Додела лиценци у систему узајамне каталогизације Србије COBISS.SR“ (Q98785010)

instance of (P21) academic journal article (Q18918145);

author P(50) Наташа Дакић (Q99281474).

Може се видети да је рад представљен идентификатором Q98785010, да је примерак (P21) класе академских чланака (Q18918145) и да има аутора Q99281474.

Већ поменути алат *OpenRefine*, који је иницијално развио Google, и алат *QuickStatements* који је развио члан тима Википодатака, Magnus Manske се често користе заједно и може се рећи да се међусобно допуњују. *QuickStatements* користи текстуални TSV или CSV формат који се ефикасно генерише алатом *OpenRefine*. Разлика између ова два алата се огледа и у грануларности трансакција, јер *OpenRefine* уноси измене у једном кораку, па разрешавање грешака насталих при уносу може довести до дуплирања података, док *QuickStatements* појединачно уноси сваку ставку и омогућава боље праћење целокупног процеса. Примери добре праксе упућују да се *OpenRefine* користи за припрему уноса у базу Википодатака, док се физички унос RDF тројки ради коришћењем алата *QuickStatements*.

Први корак је свакако припрема података у виду CSV датотеке, потом креирање *OpenRefine* пројекта и учитавање припремљених података. Даље следи препознавање постојећих ставки у Википодацима – неопходан корак који треба да омогући повезивање садржаја датотеке са идентификаторима (*QID*) постојећих ставки и унос нових, уколико оне већ не постоје. У овој фази је ручна провера и евентуална промена неопходна. Креирање схеме скупа за унос дефинише предикате који ће повезивати субјекте и објекте у RDF тројкама и врло је важан корак. Наводимо карактеристичне примере са коментарима:

- наслов (P1476), на енглеском и српском (оба писма, ћирилицом и латиницом због претраге);
- главна тема стваралачког дела (P921), кључне речи, при чему су оне које су већ постојале повезане, а нове су додате као примерци са називима ставки на српском и енглеском;
- издавач (P123);
- језик дела или имена (P407);
- датум издавања (P577), представљен само годином;
- број страна (P1104), том (P478), публикација (P433);
- објављено у (P1433) Инфотека (Q25460443);

- лиценца (P275);
- пун текст доступан на (P953).

Имајући у виду да се својства ређе додају, препорука је да се истраже слична својства и својства сличних ресурса, пре него се донесе одлука за њихово додавање. Уз то, посебну пажњу треба посветити и ограничењима својстава која се могу видети у сугестијама при уносу. Имајући у виду здружени рад дистрибуираних корисника, неминовно је да се понекад могу наћи дупликати Википодатака, а решење за те ситуације је опција спајања дупликата или елиминисања, о чему се може више видети на одговарајућој страници.

Након иницијалног уноса података, даљи унос је настављен после сваког публикованог броја *Инфотеке*, тако да се сада може пронаћи 138 радова из *Инфотеке*. Креиран је HTML који интегрише сервис за постављање упита *Wikidata Query Service* са *Библиом*. Написани су упити који добављају табеле последњих објављених радова, заступљености кључних речи у радовима, слике аутора, табелу профила аутора, граф коауторства, дистрибуције аутора по полу, и сл. Подаци о ауторима обухватају елементарне податке које свакако у наредном периоду треба допунити новим садржајима: институцијом у којој раде, областима истраживања, референцама ка референтним истраживачким базама и сл.

Као пример наводимо једноставан упит који приказује списак радова из Инфотеке, број, волумен и датум публикавања.

```
SELECT ?paper ?paperLabel ?vol ?publication ?publication_date
WHERE {
  ?paper wdt:P1433 wd:Q25460443;
         wdt:P577 ?publication_date;
         wdt:P478 ?vol;
  SERVICE wikibase:label {bd:serviceParam
                           wikibase:language "en,sr".}
}
ORDER BY DESC(?vol )
```

Илустрација претходног упита у радном окружењу *Wikidata Query Service* се може видети на слици 2, где се горњи део панела користи за креирање упита, а у доњем се приказују резултати, при чему се начин приказа може бирати у зависности од типа упита (табела, графички приказ, мрежа, приказ на временској линији, мапа и сл.).

The screenshot shows the Wikidata Query Service interface. On the left, there are filters for 'date of publication' (set to 'today') and 'language' (set to 'Serbian'). The main area displays a SPARQL query:


```

1 SELECT DISTINCT ?paper ?paperLabel ?vol ?publication ?publication_date
2 WHERE {
3   ?paper wdt:P433 wdt:Q23460443;
4   wdt:P577 ?publication_date;
5   wdt:P476 ?vol;
6   wdt:P433 ?publication.
7   RESOLVE wdt:wikiLabel ( ?id serviceParam wdt:wikiLabelLanguage "sr" ) .
8 }
9 ORDER BY DESC(?vol)
10 LIMIT 1000
11
  
```

 Below the query, the results are shown in a table with columns: paper, paperLabel, vol, publication, and publication_date. The table contains 6 rows of data, each representing a paper with its label, volume, and publication date (all dates are 1. maj 2008.).

paper	paperLabel	vol	publication	publication_date
wd:Q99231397	Accessing Scientific Information in Serbia: Six Years Experience	9	1-2	1. maj 2008.
wd:Q99231399	A Suffix Subsumption-based Approach to Building Stemmers and Lemmatizers for Highly Inflectional Languages with Sparse Resources	9	1-2	1. maj 2008.
wd:Q99231400	From the History of the Library and Information Science Department of the Faculty of Philology of the University of Belgrade	9	1-2	1. maj 2008.
wd:Q99231401	Cooperative Work in Further Development of Serbian Wordnet	9	1-2	1. maj 2008.
wd:Q99231402	Software tools for Serbian lexical resources	9	1-2	1. maj 2008.
wd:Q99231403	10th Interfending and Document Supply Conference: Resource Sharing for the Future - Building Blocks to Success (Singapore, 29th-31st October 2007)	9	1-2	1. maj 2008.

Слика 2. Окружење *Wikidata Query Service* за рад са упитима

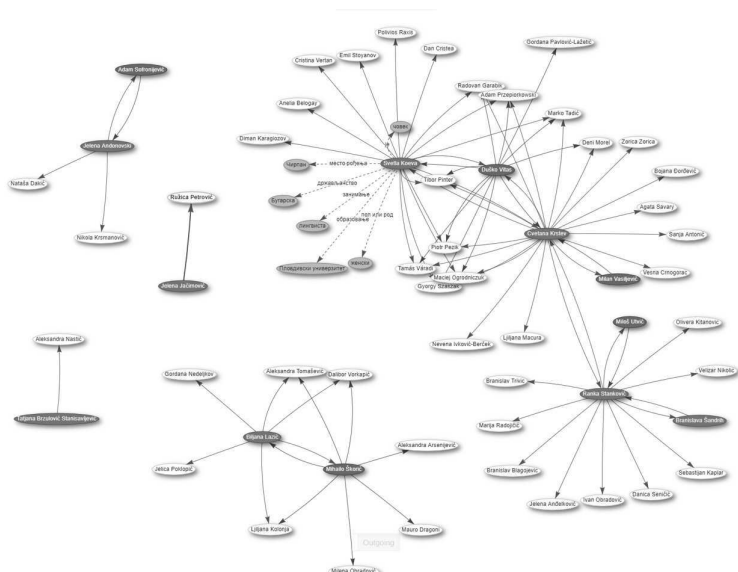
На слици 3 се може видети део графа коауторства¹⁸ који прибавља Википодатке путем сервиса за претрагу *Wikidata Query Service*.

4. Закључак и даљи планови

Позитивна искуства у раду са Википодацима *Инфотеке* су искоришћена и за унос података о романима и о ауторима романа у Википодатке из вишејезичне колекције ELTEC (European Literary Text Collection) чија ће се једна потколекција састојати од 100 српских романа из периода 1840-1920 коју у оквиру Cost акције: CA16204 – *Distant Reading for European Literary History* развијају чланови Друштва *JePTeX* под руководством Цветане Крстев и Ранке Станковић. Припремљен је скуп метаподатака за романе који су до сада дигитализовани и обрађени према захтевима Акције. Рад на Википодацима видимо као континуалну активност, при чему ће се посебна пажња обратити на повезане отворене податке из области лингвистике – LLOD и њихове примене. Морамо

18. Граф коауторства је доступан, може се приступити самом SPARQL упиту или променити опција приказа да буде табеларан, на временској скали, графички, и слично.

свакако бити свесни проблема и ограничења Википодатака и других отворених повезаних података, како бисмо истражили могућности њиховог превазилажења или бар ублажавања.



Слика 3. Граф коаутора на радовима у *Инфотеци*

Википодаци су заиста огромна база знања, која је 1) доступна свима – за читање информација, постављање упита, уређивање и унапређивање; 2) отворена – вишеструка употреба обезбеђена је бесплатном Creative Commons CC0 лиценцом, која даје пуну слободу за коришћење података; 3) вишејезична – ентитети се могу именovati и описивати на било ком природном језику. Ова три кључна својства су главни покретачи за бројне апликације, што верујемо да ће још више инспирисати вики заједницу да се овом ресурсу посвети више пажње. Такозвани мали језици, какав је српски, треба да користе све начине да се изборе за своје место у дигиталном простору, па активности на овом пројекту и сродним пројектима и иницијативама видимо као скроман допринос очувању српског језика у дигиталном добу.

Литература

- Andonovski, Jelena, Branislava Šandrih, and Olivera Kitanović. 2019. “Bilingual lexical extraction based on word alignment for improving corpus search.” *The Electronic Library*.
- Krstev, Cvetana, Jelena Jaćimović, Branislava Šandrih, and Ranka Stanković. 2019. “Analysis of the first Serbian Literature Corpus of the Late 19th and Early 20th century with the TXM platform.” In *Book of abstracts of DH_BUDAPEST_2019*, 36–37.
- Nielsen, Finn Årup, Daniel Mietchen, and Egon Willighagen. 2017. “Scholia, scientometrics and Wikidata.” In *European Semantic Web Conference*, 237–259. Springer.
- Popović, Aleksandra, Milica Ševkušić, and Đorđe Stakić. 2015. “Biblioteke i Vikipedija zajedno na vebu: slobodno znanje za sve.” *Digitalna humanistika: tematski zbornik u dve knjige, knj. 1*, 151–161.
- Shah, Urvi, Tim Finin, Anupam Joshi, R Scott Cost, and James Matfield. 2002. “Information retrieval on the semantic web.” In *Proceedings of the eleventh international conference on Information and knowledge management*, 461–468.
- Stakić, Đorđe. 2009. “Wiki Technology - Origin - Development and Importance.” *INFOtheca-Journal of Informatics & Librarianship* 10 (1-2): 69–78.
- Stanković, Ranka, Cvetana Krstev, Biljana Lazić, and Dalibor Vorkapić. 2015. “A bilingual digital library for academic and entrepreneurial knowledge management.” In *Proceeding of 10th International Forum on Knowledge Asset Dynamics-IFKAD*, 1764–1777.
- Андоновски, Јелена. 2020. “Мрежа отворених података и језички ресурси у процесу изградње српско-немачког литерарног корпуса.” PhD diss., Универзитет у Београду, Филолошки факултет, јануар.

Мултимедијални пројекат *Њих двоје*

УДК 004.55:378.147]:02(497.11)

САЖЕТАК: Тема мултимедијалног пројекта *Њих двоје* је представљање романа српских писаца у којима су ликови у романтичним везама. На пројекту су учествовали студенти завршне године Катедре за библиотекарство и информатику Филолошког факултета Универзитета у Београду. Пројекат је реализован у склопу предмета Мултимедијални документи академске 2019/20. године. Свако од студената је за један изабрани роман направио веб-страницу (помоћу HTML језика) на којој се налазе подаци о писцу и роману. Осим тога направљена је веб-страница која садржи све радове на једном месту, као и списак студената који су учествовали у изради документа.

КЉУЧНЕ РЕЧИ: мултимедијални документи, књижевност, српски романи, српски писци, љубавни романи, књижевне награде

РАД ПРИМЉЕН: 4. јуни 2021.

РАД ПРИХВАЋЕН: 30. август 2021.

Катарина Главоњић
katarina.26.g@gmail.com

Магдалена Лукић
magdalenalukic383@gmail.com

Јован Јанковић
dzovybrat@gmail.com

Стефан Јоксимовић
joksimovic.calisthenics@gmail.com

*Универзитет у Београду
Филолошки факултет
Београд, Србија*

1. Увод — О књижевности и читању

Књижевност се служи језиком као сликарство бојом, а музика тоновима, али језик није природна грађа, већ је и сам изузетно сложена духовна творевина. (Solar 2005) Када се говори конкретније о књижевности помену се књижевна дела, па и читање као важан елемент човековог живота. „Читање је често упоређивано са путовањем. (...) Као што путник види више него онај ко седи на кућном прагу, свашта искуси и неочекиване сусрете доживи, читалац открива незнане светове и живот других људи. (...) Светови песничке маште немају граница ни у простору ни у времену, а број им се увећава, стално.“ (Тартаља 1998)

Данас су књига и читање све више изложени конкуренцији нових средстава акустичке, визуелне и мултимедијалне комуникације. Тако постоје и многи романи који имају своје позоришне адаптације или су екранизовани. И у мултимедијалном пројекту *Њих двоје* могу се наћи романи који су адаптирани у позоришне представе али и они према којима су снимљени филмови и серије.

2. О роману као књижевној врсти

Роман је најпопуларнија и највише објављивана књижевна врста данас. Од 13. века појављује се *витешки роман*, а називом *le roman breton* означавана су дела (у прози и стиху) келтског и бретонског порекла у којима су описане љубавне и ратничке пустоловине хришћанских витезова. (Живковић 1992) Романи су сматрани за измишљене љубавне приче. Антички и витешки романи највише *дугују* старим еповима и бајкама, док се новији европски роман ослања на традицију шалјиве приче, анегдоте и новеле. (Тартаља 1998) У 18. веку роман је постао најзаступљенија књижевна врста. Временом је постао прича о приватом свету и о појединачним личностима, у једном реалнијем и интимнијем тону. Реакција појединца на дејство спољашњег света чини кључан део за развијања радње у роману. Од 18. до 20. века карактеризација појединих ликова у роману добија све шири значај, уопштавајући битне карактеристике једне средине и епохе. Роман приказује живот једног друштва служећи се историјом судбине појединаца, а модерног романсијера је све више почео да интересује човек сам. (Живковић 1992)

2.1 Врсте романа

Постоји доста различитих типова и врста романа. Њих можемо да разврстамо према разним критеријумима. Често се класификују према теми (историјски, авантуристички, љубавни, друштвени, тј. социјални, психолошки итд.), начину обраде теме (хумористички, сатирични, сентиментални итд.), облику у ком су (путописни, роман-есеј, роман у наставцима итд.), правцу и епохи у којој су настали (средњовековни, постмодернистички итд.). Романи могу да се разврстају и према елементима који доминирају у самом роману. Тако роман може бити и: роман лика или карактера, породични роман (који се развио из романа карактера, а у коме се углавном прате чланови једне породице),

роман простора, роман времена, као и роман времена и простора и тако даље. (Ђорђевић and Лучић 2008)

3. Укратко о српској књижевности и истакнутим писцима од 19. до 21. века

На настанак и развој српске књижевности утицали су многи српски, али и страни писци. Навешћемо само неке српске писце који су значајни за развој српске књижевности, а од којих је већина и део овог мултимедијалног пројекта.

Од свих модерних књижевних врста роман се најраније појављује у нашој књижевности. Најпре се јављају преводи, а затим, почетком 19. века, с Атанасијем Стојковићем (1773–1832) и Милованом Видаковићем (1780–1841) и оригинални романи. Атанасије Стојковић је написао дидактичко-идилични роман *Аристид и Наталија* (1801), којим започиње ова врста у нашој књижевности. Творцем и једним од најважнијих представника српског реалистичког романа у 19. веку сматра се Јаков Игњатовић (1822–1889). Од значајних романописаца 19. века треба споменути и Светолика Ранковића (1863–1899) — творца српског психолошког романа. Он је иза себе оставио неколико дела, међу којима је и значајан роман *Сеоска учитељица* (1899), који је један од првих романа где је главни лик жена. Крајем 19. века у српској реалистичкој прози увелико много се писало о селу — о пропадању села или пак његовим идиличним сликама, али и о обичајима и односима међу људима.

Велики преокрет, како у европској тако и у српској књижевности, десио се на крају 19. и почетком 20. века. У то време српска књижевност добија сва основна обележја модерне националне књижевности. Тада у српској књижевности започиње излазак из дотадашњег *сеоског* реализма. У прозне писце који су ишли ка осавремењивању прозног израза убрајају се Борисав Станковић, Петар Кочић, Милутин Ускоковић, Вељко Петровић, Исидора Секулић и други. Борисав Станковић (1876–1927) је најзначајнији приповедач у модерној фази реализма. С њим почиње развој српског модераног романа и приповетке, а његов роман *Нечиста крв* (1910) спада у најбоље романе српске књижевности.

Двадесети век је у књижевности више него раније афирмисао женско стваралаштво. Уз Исидору Секулић (1877–1958) значајне су Јелена Димитријевић (1862–1945) и Милица Јанковић (1881–1945). Може се

рећи да роман *Бакон Богородичине цркве* (1919) Исидоре Секулић и роман *Нове* (1912) Јелене Димитријевић спадају међу најбоље романе које су писале жене на почетку 20. века, у српској књижевности. Милица Јаковљевић (1887–1952), познатија под псеудонимом Мир-Јам, је књижевница чији су љубавни романи и приче обележиле домаћу књижевност између два светска рата. Она је у Краљевини Југославији била најпопуларнија књижевница и њена дела су била најчитанија.¹ А и данас се може рећи да су њени романи, али и само њено име, тј. псеудоним, веома познати српској јавности, нарочито јер су по њеним романима снимљене телевизијске серије које су изузетно гледане.

Експанзија романа, као књижевне врсте, започиње почетком педесетих година двадесетог века, а траје све до данас. Добрица Ћосић (1921–2014) је писац који је обележио послератни период. Написао је роман *Корени* (1954) који је постао темељ модерног српског послератног романа и са којим су дошла и велика признања. Након ових, а и многих других, српских писаца долазе нови и настављају традицију писања романа.

Податке о историјском развоју српских романа који су сврстани у љубавне нисмо могли да нађемо, јер је такве романе тешко наћи у историји српске књижевности. Иако постоји доста српских романа где је битан елемент и љубавни мотив они нису дефинисани као љубавни. Најчешће су дефинисани само као друштвени, историјски или романи личности итд., или спој тих жанрова. Данас су љубавни романи доста популарни и имају велике тираже, иако их већина људи сматра за мање вредна књижевна дела. Међу најпопуларнијим савременим српским књижевницама љубавних романа издвајају се Мирјана Бобић Мојсиловић (1959–), Јелена Бачић Алимпић (1969–) и Весна Дедић (1967–).²

Писци који су споменути овде, а и други који су обухваћени овим пројектом, су познати српској јавности (неки ширем, а неки ужем кругу). Већини је заједничко то да су за своја дела добили награде и да су њихова дела преведена на неки од страних језика. У част делу појединих, постоје и награде назване по њима. Ту се убрајају: Борисав Станковић, Исидора Секулић, Меша Селимовић, Милош Црњански, Момо Капор итд. Неке од награда које су писци добили за романе обухваћене овим пројектом су:

1. Преузето са Википедије (приступљено 10.8.2021.)

2. Преузето са Википедије (приступљено 10.8.2021.)

- Награда Српске књижевне задруге за доживотно дело — роман *Дошљаци* Милутина Ускоковића (1884–1915) (као најбољи рукопис 1909. године);
- НИН-ова награда³ коју је први освојио роман *Корени* Добрице Ћосића 1954. год., *Роман о Лондону* Милоша Црњанског (1893–1977), 1971. год., *Судбина и коментари* Радослава Петковића (1953–), 1993. год.;
- Награда за најчитанију књигу у библиотекама Србије⁴ — *Роман о Лондону* Милоша Црњанског 1973. год., *Доротеј* Добрила Ненадића (1940–), 1978. год.

Када је реч о романима женских писаца, јављају се и друге награде. Рукопис првог романа Гроздане Олујић (1934–2019) — *Излет у небо* је 1957. године добио награду на конкурс Сарајевске Народне просвете. Њен други роман — *Гласам за љубав* (1963) добио је награду загребачког Телеграма за најбољи кратки роман. Марија Јовановић (1959–) је за роман *Сплеткарење са сопственом душом* (2000) награђена Женским пером⁵. Вида Огњеновић (1941–) је за роман *Прелубници* (2006) примила награду за најбољу књигу у јавним библиотекама Србије⁶. А Јелена Бачић Алимпић је за свој романа *Рингишпил* (2010) награђена Златним Хит либером⁷.

4. О мултимедијалном пројекту *Њих двоје*

Главна тема пројекта је представљање романа српских писаца у којима постоје ликови који су у романтичним везама. Уз то су представљени и писци и романи. Ти романи су по типу разноврсни — нису сви чисто љубавни, али у себи имају љубавни мотив као један од главних елемената романа.

3. Престижна српска, а раније југословенска, књижевна награда која се сваке године додељује за најбољи нови српски (раније југословенски) роман.

4. Додељивала ју је Народна библиотека Србије од 1973–2004. године, уз прекиде.

5. Додељивао га је женски часопис Политика Базар, за најбољи роман или збирку прича женске списатељице.

6. Додељивала ју је Народна библиотека Србије од 2005–2013. године.

7. Признање ауторима и издавачима најтраженијих књига у текућој години, које је додељивала Редакција за културу и уметност РТС-а и емисија Хит либрис.

Љубавних романа, и оних који су уврштени у два жанра од којих је један љубавни, има око 10^8 , док је више оних романа који спадају у романе личности, друштвене и историјске романе. У пројекту се могу пронаћи највише романи из 20. века али има и неколико из 19. и 21. века. Укупно их је 25, колико је било и учесника овог пројекта.

Романи су представљени у форми веб-страница које су направили студенти завршне године Катедре за библиотекарство и информатику Филолошког факултета у Београду, академске 2019/2020. године. Помоћ студентима пружиле су проф. др Цветана Крстев и доц. др Брнислава Шандрих Тодоровић, које су биле менторке на пројекту. Овај пројекат је саставни део предмета Мултимедијални документи који је део наставног програма у четвртој години основних студија на Катедри за библиотекарство и информатику.

5. Романи и писци који су обухваћени овим пројектом

Романи које су студенти изабрали да прикажу су: *Аристид и Наталија* (1801) Атанасија Стојковића, *Чудан свет* (1869) Јакова Игњатовића, *Сеоска учитељица* (1898) Светолика Ранковића, *Дошљаци* (1910) Милутина Ускоковића, *Нечиста крв* (1910) Борисава Станковића, *Нове* (1912) Јелене Димитријевић, *Чедомир Илић* (1914) Милутина Ускоковића, *Вакон Богородичине цркве* (1919) Исидоре Секулић, *Плава госпођа* (1924) Милице Јанковић, *Покошено поље* (1934) Бранимира Ћосића, *Српска трилогија* (1937) Стевана Јаковљевића, *Песма* (1952) Оскара Давича, *Корени* (1954) Добрице Ћосића, *Излет у небо* (1958) и *Гласам за љубав* (1963) Гроздане Олујић, *Тврђава* (1970) Меше Селимовића, *Роман о Лондону* (1971) Милоша Црњанског, *Доротеј* (1977) Добрила Ненадића, *Уна* (1981) Моме Капора, *Цвет лепе на Балкану* (1991) Гордане Куић, *Судбина и коментари* (1993) Радослава Петковића, *Сплеткарење са сопственом душом* (2000) Марије Јовановић, *Прељубници* (2006) Виде Огњеновић, *Рингиштил* (2010) Јелене Бачић Алимпић и *Никад нисам* (2019) Весне Дедић.

Од романа који су били познати већини (пре почетка рада на овом пројекту) издавају се: *Нечиста крв*, *Корени*, *Тврђава*, *Цвет лепе на*

8. Због недостатка јасних података о жанру овај број није прецизан.

Балкану и Дошљаџи. Романи за које већи број студената раније није знао су: *Сплеткарење са сопственом душом*, *Прељубници*, *Аристид и Наталија*, *Нове*, *Сеоска учитељица*, *Плава госпођа*, *Чудан свет*, *Доротеј*, *Судбина и коментари*. Од писаца за које је већина знала издвајају се: Борисав Станковић, Исидора Секулић, Оскар Давичо, Добрица Ћосић, Меша Селимовић, Милош Црњански, Момо Капор, Јелена Бачић Алимпић и Весна Дедић. Студентима су скоро непознати били писци: Марија Јовановић, Милица Јанковић, Добрило Ненадић, Атанасије Стојковић, Светолик Ранковић и Јелена Димитријевић.

5.1 Подела одабраних романа по жанру

Што се тиче типолошке тј. жандровске поделе наведених романа, треба нагласити да је податке о томе теже наћи. За романе као што су *Излет у небо* и *Роман о Лондону* нису пронађене жанровске поделе, а за роман *Доротеј* се наводи да је на граници жанрова — да у себи садржи историјске, љубавне и психолошке мотиве⁹. Од свих изабраних романа, романи који су сврстани у љубавне су: *Плава госпођа*, *Гласам за љубав*, *Уна*, *Сплеткарење са сопственом душом*, *Прељубници*, *Рингишпил* и *Никад нисам*. А поред тих и романи: *Дошљаџи*, *Чедомир Илић* и *Бакон Богородичине цркве*, који спадају и у друге жанрове — оба Ускоковићева романа спадају и у друштвене, а роман Исидоре Секулић спада и у религиозан роман. Остали романи су сврстани у различите жанрове (углавном у друштвени, историјски, психолошки и у роман личности).

5.2 Романи који имају драмске адаптације или су екранизовани

Један од значајних елемената пројекта је и податак о адаптацијама романа. Што се тиче драматизације и екранизације ових романа њих седам (од 25) нема своју позоришну ни телевизијску адаптацију.¹⁰ Ту спадају романи: *Аристид и Наталија*, *Нове*, *Плава госпођа*, *Сплеткарење са сопственом душом*, *Прељубници*, *Рингишпил* и *Никад нисам*.

Четрнаест романа је адаптирано за позоришно извођење. То су: *Чудан свет*, *Сеоска учитељица*, *Дошљаџи*, *Нечиста крв*, *Чедомир*

9. Податак преузет са сајта онлајн књижаре Laguna (приступљено 15.8.2021.)

10. Подаци о томе нису пронађени путем претраге интернета.

Илић, Ђакон Богородичине цркве, Покошено поље, Српска трилогија, Песма, Корени, Излет у небо, Тврђава, Роман о Лондону и Судбина и коментари. Међу тим представама има и оних са другачијим називом од оригиналног наслова романа. Ту се убрајају представе тј. романи: *Врела крв (Нечиста крв), Силе (Покошено поље), Хроми идеали (Чедомир Илић), Ђакон (Ђакон Богородичине цркве)*; по мотивима овог романа и приповетке *Амбиције, дим Исидоре Секулић*) и *На леђима језа (Српска трилогија)*.¹¹

По мотивима осам романа снимљени су филмови: *Песма* (1961), *Излет у небо* (1962. под називом *Чудна девојка*), *Гласам за љубав* (1965), *Дошљаци* (1969), *Покошено поље* (1980. под називом *Београдска разгледница 1920*), *Доротеј* (1981), *Уна* (1984) и *Нечиста крв* (1996. и 2021.). Романи по чијим мотивима су снимљене телевизијске серије су: *Чедомир Илић* (1971), *Песма* (1975), *Роман о Лондону* (1988), *Цвет лепе на Балкану* (2011-2012) и *Корени* (2018).¹²

6. Задужења студената на пројекту

Пре почетка израде самог мултимедијалног документа било је потребно урадити неколико задатака. Први задатак је био да свако од студената пронађе књиге које би одговарале теми пројекта. Сви су, заједно са професоркама, дали предлог једне или више књига, а онда је, на наредном часу, свако изабрао једну коју ће да прочита код куће и после, на неком од часова, свима да је представи у виду презентације (користећи програм MS PowerPoint). Презентације су садржале основне податке о писцима и књигама — име и презиме писца, наслов дела, годину првог издања, биографију писца, листу ликова и опис теме романа. Поједине су садржале и неке занимљивости из романа или податке о томе да ли су ти романи екранизовани или имају позоришне адаптације. Наредни и главни задатак је био да се на основу тога,

11. Подаци су преузети са сајта Музеја позоришне уметности Србије (приступљено 10.8.2021.)

12. Подаци о екранизацијама углавном су преузети са Википедије и сајта IMDb.

код куће, направе веб-странице коришћењем језика HTML,¹³ уз CSS¹⁴ за визуелно обликовање и улепшавање. Свако је за изабрани роман самостално (по својој жељи и знању) осмислио и израдио веб-страницу.

Поред тих задатака додељени су и групни задаци — израда веб-страница за цео пројекат (осмишљавање дизајна и израда HTML кода) и избор музике, тј. песме која ће да оживи пројекат. За те групне задатке оформљене су две групе студената. Групу која је била задужена за музику су чинили студенти: Марија Деспић, Ана Ивановић, Аница Јовановић, Софија Хебар и Никола Шибалић. Другу групу која је била задужена за осмишљавање дизајна и саму техничку израду чинили су студенти: Катарина Главоњић, Магдалена Лукић, Јован Јанковић и Стефан Јоксимовић. Група задужена за музику је најпре направила списак песама са *Јутјуба* (YouTube), након чега је уследило гласање за песму која ће да се користи на сајту. Одабрана је адаптација композиције *Љубавна прича* (Love Story) Ричарда Клејдермана (Richard Clayderman) из истоименог филма.

За бољу организацију задатака, поред имејл преписке, користио се и *Гугл диск* (Google Drive). Ту су били постављени сви потребни документи и информације о задацима, које су биле доступне свим студентима у сваком тренутку. За списак романа и задужења студената направљена је табела (у бесплатном онлајн програму Google Sheets), која се временом допуњавала. Такође ту су направљени и фолдери где су студенти постављали своје радове (презентације, HTML документе и пропратне елементе које су коришћене за израду веб-страница, као што су фотографије и сл.).

6.1 Прикупљање података

Проналажење романа који одговарају теми пројекта (оних чији садржај обухвата причу људи између којих постоји романтична веза) није био много тежак задатак. На почетку су постојале мале недоумице око саме теме тј. критеријума за одабир романа и писаца, јер студенти нису били довољно упознати са таквим романима. Али релативно брзо,

13. HTML (HyperText Markup Language) – Језик за означавање хипертекста, односно описни језик специјално намењен опису веб-страница. Преузето са Википедије (HTML) (приступљено 27.5.2021.)

14. CSS (Cascading Style Sheets) — Каскадни стилски листови, језик форматирања помоћу ког се дефинише изглед елемената веб-странице. Преузето са Википедије (приступљено 27.5.2021.)

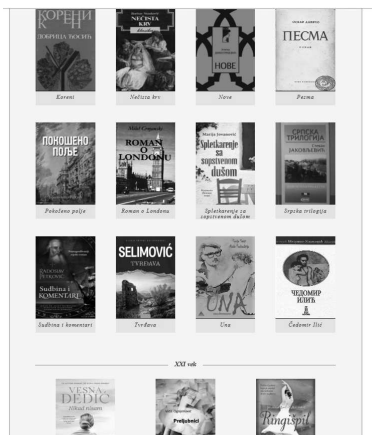
уз помоћ професорки, направљен је списак одговарајућих романа. На часовима се дискутовало о предлозима и тако је направљен почетни списак, кога су студенти касније допунили. Сви су имали задатак да пронађу бар један роман и да својим предлозима допуне табелу (тако што би додали наслов, име аутора и годину првог издања романа). Тако је формирана коначна листа романа.

Разлози за избор баш ових романа били су различити — неко је већ раније чуо за роман који одговара теми или га је прочитао, неко је тражио препоруке (углавном од чланова своје породице), а неки студенти су самостално, претрагом интернета дошли до изабраног романа, док су остали изабрали неки од понуђених са списка. Несумњиво су и значај и популарност већине писаца и романа допринели томе да се студенти одлуче да прикажу баш та дела српске књижевности.

Након коначног избора романа студенти су их прочитали, а потом прикупили податке о писцима и романима. Један од главних извора свакако су били изабрани романи, које су студенти анализирали и представили (фокусирајући се на тему овог пројекта). Поред саме садржине романа, коришћени су и подаци са корица књига, као и из предговора и поговора. Други извори информација били су углавном Википедија, сајтови писаца и електронске новине (са разним чланцима о писцима и романима, а који садрже и податке о представама, филмовима и сл. као и интервјуе са писцима) али и разни други сајтови (претежно због проналажења фотографија, видео снимака и других детаља и занимљивости).

7. Веб-странице пројекта

Мултимедијални пројекат *Њих двоје* чине три основне веб-странице. Прва страница пројекта садржи све романе груписане према периоду настанка, од 19. до 21. века (слика 1). Ту су за приказ романа коришћене фотографије корица књига, а испод сваке стоји наслов романа. Кликом на те фотографије отварају се странице које су студенти направили (слике 2 и 3 приказује део HTML и CSS кода једне веб-странице). Свака од тих страница се изгледом и садржином разликује од осталих. Друга страница садржи фотографије свих студената који су учествовали у пројекту док трећа страница садржи податке о самом пројекту (Слика 4) — са којим циљем је пројекат осмишљен, ко су аутори пројекта, шта садрже веб-странице, а на крају су поменути и студенти који су радили на изради сајта (осмишљавању дизајна



Слика 1. Приказ приступне странице где се налазе сви романи

и изради главних страница). Допуна овим страницама су и линкови ка сајту Филолошког факултета у Београду и страници Катедре за библиотекарство и информатику, а такође и линкови ка сајтовима проф. др Цветане Крстев и доц. др Браниславе Шандрих Тодоровић. Саставни део сајта чини и аудио-плејер са адаптацијом композиције *Љубавна прича* Ричарда Клејдермана кога је тим за техничку израду (Јован Јанковић и Стефан Јоксимовић) поставио на сајт.

Сама израда (појединачних и главних) веб-страница већини није представљала тежак задатак. Студенти су имали мале потешкоће при раду јер је било потребно да се подсети HTML-а, а неки су морали самостално да науче нешто више о њему (од онога што је предавано на часовима), јер је њихова замисао дизајна страница захтевала веће знање. Због свега тога, за израду радова је требало мало више времена и труда али на крају су сви студенти успешно и у задатом року завршили радове (и упркос избијању пандемије Ковид 19).

8. Садржински слични пројекти

Мултимедијални документи су саставни део наставног програма од школске 2009/2010. Ако се осврнемо уназад приметимо да постоји веза између овогодишњег и неких ранијих мултимедијалних докумената. Пошто је у пројекту *Њих двоје* акценат на романима српских писаца, мултимедијални пројекат ранијих генерација студената



Слика 4. Приказ треће странице са детаљима о самом пројекту

који је најсличнији овом је пројекат *Књижевност на филму* (Меглић 2019), а затим и пројекат *Поп Тира и Поп Спира* (Коврлија, Тасић, and Топаловић 2012). Ови радови обухватају романе српских писаца и то оне који су екранизовани, а екранизација романа је један од додатних елемената у пројекту *Њих двоје*. Поред њих, по теми сличан је и пројекат *Пут око света за 80 дана* (Перић, Гогић, and Николић 2014) јер и он обухвата роман који је екранизован.

9. Закључак

Пре почетка рада на пројекту *Њих двоје*, студентима су били познати многи писци и романи који су се нашли на њиховом списку романа погодних за овај пројекат. Сви су углавном лако дошли до потребних података (информација о писцу, роману, фотографија и сл.), но мањи проблеми су се јавили приликом израде веб-страница јер је требало да се студенти подсети HTML-а, који су учили на другој години студија. Мали проблем неким је било и време потребно за израду веб-страница, али је на крају све завршено без већих проблема и у предвиђеном року (иако је пандемија Ковид 19 донекле изменила првобитни план програма рада).

Овај пројекат студентима није био превише захтеван, имали су прилику да искористе и унапреде своја знања стечена на факултету али и да буду креативни свако на њему својствен начин. Открили су и прочитали нове књиге, сазнали нешто више о писцима, фиктивним ликовима или стварним људима, као и друштву, месту, периоду у ком

су они живели итд. Потом су део тога испричали на часовима, а затим, након неког времена, направили веб-странице за те романе. На крају су их све објединили у један заједнички пројекат, доступан свима на интернету.¹⁵

Пројекат је објединио истраживачки рад (проналажење информација на интернету), знање HTML-а, CSS-а и PowerPoint програма. Све то су студенти могли да науче, или побољшају, током студија на Катедри за библиотекарство и информатику. Учење PowerPoint-а је део предмета Информатички практикум 1, који се одржава у првој години основних студија при Катедри, а HTML и CSS су део предмета Дигитални текст 2, који се одржава у другој години. Што се тиче истраживачког рада он се кроз целе студије унапређивао кроз разне задатке на многим предметима, а у оквиру овог предмета достигао је још виши ниво.

Изјава захвалности

Желимо да упутимо посебну захвалност професоркама др Цветани Крстев и др Бранислави Шандрих Тодоровић које су увек биле доступне за сарадњу и помоћ и које су нас успешно водиле кроз овај пројекат из предмета Мултимедијални документи, академске 2019/2020. године.

Литература

Solar, Milivoj. 2005. *Teorija književnosti*. Zagreb : Školska knjiga.

Ђорђевић, Часлав, and Предраг Лучић. 2008. *Књижевност и српски језик : приручник за ученике четвртог разреда гимназије и средњих стручних школа*. 73–74. Нови Сад: Д. Ђорђевић, Венцловић.

Живковић, Драгиша. 1992. *Теорија књижевности са теоријом писмености : стилистика, књижевни родови и врсте, књижевни периоди и књижевни правци, о композицији и стилу писаних састава : приручник за наставнике и ученике*. Београд : Завод за уџбенике и наставна средства.

15. На сајту доц. др Браниславе Шандрих Тодоровић и са матичне странице проф. др Цветане Крстев.

- Коврлија, Дарја, Валентина Тасић, and Сузана Топаловић. 2012. “Израда мултимедијални документа „Поп Тира и поп Спира“, уједињено знање на заједничком студентском пројекту.” *Инфотека* 8 (2): 86–89. http://infoteka.bg.ac.rs/pdf/Srp/2012-2/INFOTHECA_XIII_2_December2012_86-89.pdf.
- Меглић, Катарина. 2019. “Израда мултимедијалног документа „Књижевност на филму“.” *Инфотека* 19 (1): 96–107. <https://infoteka.bg.ac.rs/ojs/index.php/Infoteka/article/view/160>.
- Перић, Катарина, Кристина Гогић, and Ана Николић. 2014. “Израда мултимедијалног документа „Пут око света за 80 дана“.” *Инфотека* 15 (2): 56–62. http://infoteka.bg.ac.rs/pdf/Srp/2014-2/Srp2014-2INFOTHECA_XV_2_april_56-62.pdf.
- Тартаља, Иво. 1998. *Теорија књижевности : за средње школе*. Београд : Завод за уџбенике и наставна средства.

EUROLAN 2021: Увод у повезане податке у лингвистици – онлајн школа за обуку полазника

УДК 81: 37.018.53

САЖЕТАК: Прва школа за обуку полазника коју је организовала COST акција *NexusLinguarum* одржана је од 8. до 12. фебруара 2021. године са циљем да студенти, истраживачи и стручњаци науче основе лингвистичке науке о подацима. Током обуке полазници су се упознали са широким спектром тема: од семантичког веба, RDF -а и онтологија, до моделирања и претраживања језичких података помоћу најсавременијих онтолошких модела и алата. Школа је одржана у оквиру серије летњих школа EUROLAN-а и организовало ју је виртуелно (онлајн) неколико института; Румунска академија, Истраживачки институт за вештачку интелигенцију у Букурешту и Институт за рачунарске науке, као и Универзитет „Alexandru Ioan Cuza“ у Јашију, Румунија. Школу су похађала 82 полазника.

КЉУЧНЕ РЕЧИ: наука о лингвистичким подацима, повезани подаци у лингвистици, језички подаци, EUROLAN, *NexusLinguarum*, COST акција, школа за обуку

РАД ПРИМЉЕН: 30. јун 2021.

РАД ПРИХВАЋЕН: 13. јул 2021.

Милан Дојчиновски
milan.dojchinovski@fit.cvut.cz
CTU

Праг, Чешка Република
InfAI при Универзитету
у Лајпцигу
Лајпциг, Немачка

Хулија Боске Хил
jbbosque@unizar.es
Хорхе Грасија
jogracia@unizar.es
Арагонски институт за
инжењерска истраживања
(ИЗА)
Универзитет у Сарагоси
Сарагоса, Шпанија

Ранка Станковић
ranka.stankovic@rgf.bg.ac.rs
Универзитет у Београд
Рударско-геолошки факултет
Београд, Србија

1. Увод

COST акција CA18029 *NexusLinguarum* - European network for Web-centered linguistic data science¹ „Европска мрежа за науку о

1. *NexusLinguarum*

језичким подацима оријентисаним ка вебу“ започела је активности крајем октобра 2019. Циљ акције *NexusLinguarum* је промовисање проучавања лингвистичких наука о подацима, за које је потребна изградња екосистема вишејезичних и семантички интероперабилних језичких података. Школе су једно од средстава за постизање овог циља, па је стога главни тим акције *NexusLinguarum* организовао школу „Увод у повезане лингвистичке податке“,² која је одржана од 8. до 12. фебруара 2021. Школа је имала за циљ промовисање и предавање о основама науке о лингвистичким подацима и сродним технологијама полазницима из академске и шире заједнице. Организована је под окриљем серије летњих школа EUROLAN-а основане 1993. године и обухвата теме које су посебно релевантне за области рачунарске лингвистике и обраде природних језика (ОПЈ). Циљ ове 15. школе EUROLAN-а био је да окупи научнике, наставнике и студенте лингвистике, ОПЈ-а и информационих технологија како би разговарали о принципима и најбољим праксама за представљање, објављивање и повезивање лингвистичких података и питања која чине саставне делове замишљеног вишејезичног и интероперабилног екосистема оријентисаног ка вебу. Овај прилог резимира организацију, садржај и резултате школе обуке и заснован је на извештају Д1.1 акције доступном на вебу.³

2. Програм школе

Школа је осмишљена за почетнике, као и за оне који већ имају основно знање из области обухваћених темама. Школа је пружила свеобухватан увод у методологије представљања језичких ресурса коришћењем технологија семантичког веба, заједно са средствима за екстракцију знања из језичких ресурса и његово коришћење помоћу упитних језика и механизма закључивања семантичког веба. У оквиру школе су обрађиване следеће теме:

- Семантички веб и Повезани подаци⁴ (Berners-Lee et al. 2006);
- Онтологије: RDF (Resource description framework), RDF Schema (Resource Description Framework Schema, скраћено RDFS, RDF(S), RDF-S, или RDF/S), Web Ontology Language (OWL),⁵ etc.);

2. EUROLAN

3. Извештај Д1.1

4. Увод у повезане податке и семантички веб

5. OWL

- Упитни језик SPARQL - семантички упитни језик за базе података за претраживање и руковање подацима ускладиштеним у RDF формату; item Метаподаци: DCAT (Data Catalog Vocabulary),⁶ VOID (RDF Schema речник за представљање метаподатака о RDF скуповима података, итд.);
- Трансформација и валидација RDF података (Cimiano et al. 2020);
- Лингвистички повезани подаци (Chiarcos et al. 2013);
- Lemon-OntoLex⁷ (McCrae et al. 2017; Declerck, Tiberius, and Wandl-Vogt 2017; Stanković et al. 2018)
- Генерисање лингвистички повезаних података (Cimiano et al. 2020);
- Корпуси и повезани подаци (Chiarcos 2012);
- Лингвистичке анотације (Fäth et al. 2020);
- Формат за размену у обради природних језика – NLP Interchange Format (NIF) (Hellmann et al. 2013);
- Алати и апликације лингвистичких повезаних података (Declerck et al. 2020).

Први дан почео је уводном сесијом и кратким уводом у лингвистичке повезане податке (Linguistic Linked Data, LLD), након чега су уследиле сесије посвећене уводу у повезане податке и RDF. Други дан је обухватио теме везане за онтологије, укључујући моделирање знања помоћу онтологија, језике за представљање знања OWL и SKOS,⁸ образлагање знања и практичну сесију уз употребу едитора онтологије Protégé.⁹ Трећи дан био је посвећен темама у вези са представљањем и испитивањем лексичких података са наменским сесијама о моделу Ontolex-Lemon и упитном језику SPARQL. Четврти дан је укључивао сесије које су дале преглед других лингвистичких речника и речника метаподатака, као и платформе VocBench¹⁰ (Stellato et al. 2020) за моделирање лингвистичких скупова података. У поподневним часовима организовано је виртуелно дружење на коме су учесници могли да виде видео записе лепоте румунске културе, традиције и природе. Пети дан се састојао од три паралелне сесије на различите теме:

6. Data Catalog Vocabulary (DCAT)

7. Lemon - Модел лексикона за онтологије; Лексикон модел за онтологије: Извештај заједнице од 10. маја 2016.

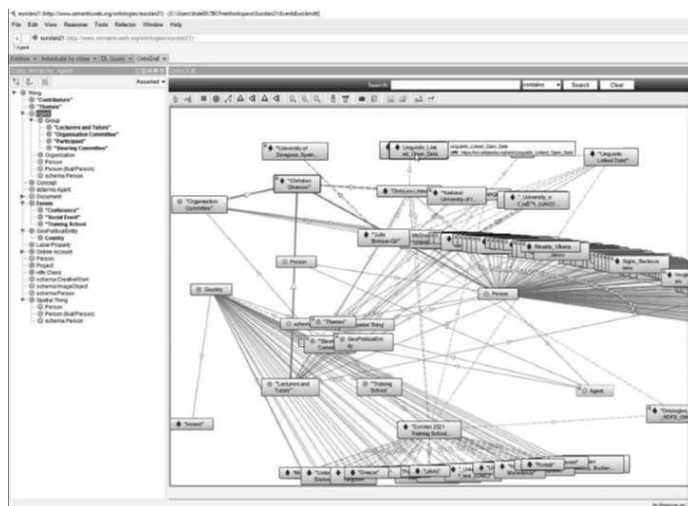
8. SKOS

9. Protégé

10. VocBench: Систем колаборативног управљања за OWL онтологије, SKOS(/XL) тезаурусе, OntoLex-lemon лексиконе и скупове података у генеричком RDF.

- (a.) Генерисање/трансформација и повезивање LLD-a,
- (b.) Анотације (NIF, веб анотације) (Hellmann et al. 2013) и
- (c.) OntoLex проширења: *vartrans* за представљање превода и варијанте термина (засновано на модулу *lemon*-a за превођење (Gracia et al. 2014), *lexicog*¹¹ - лексикографски модул (Bosque-Gil, Gracia, and Montiel-Ponsoda 2017), *FrAC*¹² - frequency, attestation and corpus Information (Chiarcos et al. 2020) – фреквенције, потрвде и информације о корпусу.

За крај, школа обуке је окончана завршном сесијом где је представљена онтологија учесника, предавача и организатора која је илустровала многе механизме представљања објашњене током недеље.



Слика 1. Онтологија учесника, предавача и организатора школе.

Свака од организованих сесија била је пропраћена практичним заједничким и самосталним вежбама. Током практичне сесије, предавачи би предложили вежбу и понудили решавање корак по корак, како би учесници могли да разумеју методологију којом се

11. The OntoLex Lemon Lexicography Module

12. FrAC – Frequency, Attestation and Corpus Information - Ontology-Lexica Community Group

долази до решења. Уз то су представили и основне технологије потребне за решавање задатка. Затим се, током сесије самосталних вежби, од учесника тражило да раде на одређеном задатку попут случајева представљених током практичне сесије, чиме би се упознали са технологијом уведеном у практично окружење. Како су ове сесије биле градиране по сложености, почевши од основних појмова, који су надограђивани тако да су последњег дана детаљно представљене конкретније теме, учесници су имали прилику да стекну чврсте темеље знања пре него што пређу на сложеније сесије. Званични програм школе доступан је на веб¹³.

Као наставак активности, Друштво за језичке ресурсе и алате JePTex¹⁴ је поставило локалну инсталацију алата VocBench¹⁵, а осим чланова друштва JePTex, користили су је и студенти и наставници програма докторских студија „Интелигентни системи“ Универзитета у Београду за предмет „Представљање знања“ и „Семантички веб“. Модул *Lemon-OntoLex FrAC* примењиван је за представљање уноса из лексикона који се користи за откривање увредљивог обраћања са потврдама из корпуса твитова са анотацијом увредљивих фраза (Jokić et al. 2021).

3. Organization

Због пандемије COVID -19 и ограничења путовања у Европи и шире, школа је одржана онлајн. Следећи традицију EUROLAN-а дугу готово три деценије, која је позната по изврности академских програма, заједно са дружењима међу професорима и студентима, поред онлајн наставе организован је и низ виртуелних активности са циљем упознавања културе и успостављања ближе интеракције. Присуство је било искључиво онлајн и бесплатно, уз захтев за претходну регистрацију.

Све сесије су биле организоване на платформи за видео конференције. За практичне сесије, било је доступно је неколико виртуелних соба у којима су учесници могли да раде на задацима у мањим групама. Да би охрабрили учеснике да постављају питања и ступају у међусобни контакт, организатори су поставили Slack¹⁶ канал,

13. Програм школе

14. Jerteh

15. Инсталација VocBench-a

16. Slack, Виртуелни центар за сарадњу

као центар за сарадњу, где су предавачи и учесници могли да разјасне све недоумице. Укупан број учесника био је 82, од тога су 52 полазника биле жене, а 30 мушкарци, укључујући и четири учесника из Србије.

За школу су осмишљене различите врсте материјала, укључујући презентације (слајдове)¹⁷ и вежбе¹⁸ пропраћене примерима кода и података.¹⁹ Сви материјали су објављени на интернету и јавно су доступни.

4. Резиме

Школа је пружила драгоцену знања и обучила многе информатичаре и лингвисте о томе како да раде и имају користи од лингвистичких повезаних података. Ово је била прва школа коју је организовала *NexusLinguarum* COST акција из низа обука чије се одржавање планира. Она је имала за циљ да послужи као увод у тему науке о лингвистичким подацима и изгради основу за публику неопходну за похађање наредних школа обуке о напреднијим темама током трајања акције. Сви материјали настали током школе су јавно доступни и заједница их може даље користити. Током завршне сесије, организатори су учесницима дали образац за прикупљање повратних информација о организационим и академским аспектима школе. Резултати показују да су хуманистичке науке, лингвистика и лексикографија имале већу заступљеност међу учесницима него рачунарство, те да је школа била добро фокусирана, тематски уравнотежена и добро организована. Теоријске сесије, подучавање и могућности учења били су веома добро оцењени. С друге стране, због виртуелног начина рада, још увек постоји простор за побољшање у практичним сесијама, организацији друштвених догађаја и могућностима за умрежавање. Стечена знања и вештине ће побољшати развој српских језичких ресурса и помоћи у објављивању више ресурса као лингвистичких повезаних података.

Захвалност

Овај рад је подржала COST акција CA18209 - *NexusLinguarum* “European Network for Web-centred Linguistic Data Science” (*NexusLinguarum* „Европска мрежа за науку о језичким подацима оријентисаним ка вебу“).

17. Презентације

18. Вежбе

19. Додатни материјали

Литература

- Berners-Lee, Tim, Yuhsin Chen, Lydia Chilton, Dan Connolly, Ruth Dhanaraj, James Hollenbach, Adam Lerer, and David Sheets. 2006. "Tabulator: Exploring and analyzing linked data on the semantic web." In *Proceedings of the 3rd international semantic web user interaction workshop*, 2006:159. Athens, Georgia.
- Bosque-Gil, Julia, Jorge Gracia, Guadalupe Aguado-de-Cea, and Elena Montiel-Ponsoda. 2015. "Applying the ontolex model to a multilingual terminological resource." In *European Semantic Web Conference*, 283–294. Springer.
- Bosque-Gil, Julia, Jorge Gracia, and Elena Montiel-Ponsoda. 2017. "Towards a Module for Lexicography in OntoLex." In *LDK Workshops*, 74–84.
- Bosque-Gil, Julia, Jorge Gracia, Elena Montiel-Ponsoda, and Asunción Gómez-Pérez. 2018. "Models to represent linguistic linked data." *Natural Language Engineering* 24 (6): 811–859.
- Chiarcos, Christian. 2012. "Interoperability of corpora and annotations." In *Linked Data in Linguistics*, 161–179. Springer.
- Chiarcos, Christian, Maxim Ionov, Jesse de Does, Katrien Depuydt, Fahad Khan, Sander Stolk, Thierry Declerck, and John Philip McCrae. 2020. "Modelling frequency and attestations for ontolex-lemon." In *Proceedings of the 2020 Globalex Workshop on Linked Lexicography*, 1–9.
- Chiarcos, Christian, John McCrae, Philipp Cimiano, and Christiane Fellbaum. 2013. "Towards open data for linguistics: Linguistic linked data." In *New Trends of Research in Ontologies and Lexical Resources*, 7–25. Springer.
- Cimiano, Philipp, Christian Chiarcos, John P McCrae, and Jorge Gracia. 2020. "Converting language resources into linked data." In *Linguistic Linked Data*, 163–180. Springer.
- Declerck, Thierry, John Philip McCrae, Matthias Hartung, Jorge Gracia, Christian Chiarcos, Elena Montiel-Ponsoda, Philipp Cimiano, Artem Revenko, Roser Sauri, Deirdre Lee, et al. 2020. "Recent developments for the linguistic linked open data infrastructure." In *Proceedings of the 12th LREC*, 5660–5667.

- Declerck, Thierry, Carole Tiberius, and Eveline Wandl-Vogt. 2017. “Encoding lexicographic data in lemon: Lessons learned.” In *Proceedings of the LDK workshops: OntoLex, TIAD and Challenges for Wordnets. CEURS*, vol. 8.
- Fäth, Christian, Christian Chiarcos, Björn Ebbrecht, and Maxim Ionov. 2020. “Fintan-flexible, integrated transformation and annotation engineering.” In *Proceedings of the 12th LREC*, 7212–7221.
- Gracia, Jorge, Elena Montiel-Ponsoda, Daniel Vila-Suero, and Guadalupe Aguado-De-Cea. 2014. “Enabling Language Resources to Expose Translations as Linked Data on the Web.” In *Proceedings of the 9th LREC*, edited by Nicoletta Calzolari (Conference Chair) et al. Reykjavik, Iceland: European Language Resources Association (ELRA), May. ISBN: 978-2-9517408-8-4.
- Hellmann, Sebastian, Jens Lehmann, Sören Auer, and Martin Brümmer. 2013. “Integrating NLP using linked data.” In *International Semantic Web Conference*, 98–113. Springer.
- Jokić, Danka, Ranka Stanković, Cvetana Krstev, and Branislava Šandrih. 2021. “A Twitter Corpus and lexicon for abusive speech detection in Serbian.” In *Proceedings of the 2021 Language, Data and Knowledge (LDK), 1-3 September in Zaragoza, Spain*.
- McCrae, John P, Julia Bosque-Gil, Jorge Gracia, Paul Buitelaar, and Philipp Cimiano. 2017. “The Ontolex-Lemon model: development and applications.” In *Proceedings of eLex 2017 conference*, 19–21.
- Stanković, Ranka, Cvetana Krstev, Biljana Lazić, and Mihailo Škorić. 2018. “Electronic dictionaries-from file system to lemon based lexical database.” In *Proceedings of the 11th LREC - W23 6th Workshop on Linked Data in Linguistics: Towards Linguistic Data Science (LDL-2018), LREC 2018, Miyazaki, Japan, May 7-12, 2018*, 48–56.
- Stellato, Armando, Manuel Fiorelli, Andrea Turbati, Tiziano Lorenzetti, Willem Van Gemert, Denis Dechandon, Christine Laaboudi-Spoiden, Anikó Gerencsér, Anne Waniart, Eugeniu Costetchi, et al. 2020. “VocBench 3: A collaborative Semantic Web editor for ontologies, thesauri and lexicons.” *Semantic Web* 11 (5): 855–881.

COBISS Meet 2020 - онлајн конференција Презентација нових могућности за све учеснике мреже COBISS.net

УДК

Ана Васиљевић
ana.ivkovic.va@gmail.com

Јелица Зељић
jelicazelj1971@gmail.com

РАД ПРИМЉЕН:
РАД ПРИХВАЋЕН:

18. март 2021
1. јул 2021

*Народна библиотека
Осечина, Србија*

Институт за информационе науке у Марибору (ИЗУМ) организовао је од 1. до 3. децембра 2020. године традиционалну COBISS конференцију под називом COBISS Meet 2020. Овог пута окупљање је обављено онлајн с обзиром на тренутну ситуацију са пандемијом Ковид-19. Конференција се одвијала на интернет платформи <https://hopin.to/events/cobiss-meet-2020>. Виртуелна сцена је била одлична прилика за дискусију о идејама и заједничким плановима за будућност, контакт са предавачима, колегама и другим учесницима догађаја. Догађај је пратило преко 600 учесника из више од 10 земаља. Централна тема била је презентација нових приступа и представљање иновација које у последње време спроводи ИЗУМ за све учеснике мреже COBISS.net.

Првог дана конференције на дневном реду били су поздравни говор, предавања и дискусија са предавачима из Словеније. У поздравном обраћању др Ален Бошњак, директор ИЗУМ-а, осврнуо се на невиђене околности у којима се нашао свет, али и на то да нас то није спречило да заједно испуњавамо своју мисију без прекида. И у таквим условима морамо омогућити нашим корисницима да у потпуности користе наше услуге. Затим су најављена предавања које ће се одржати и наглашено да ће сви имати могућност за размену мишљења са излагачима и осталим учесницима након сваке сесије, што је изванредна прилика за размену идеја и прављење заједничких планова за будућност.

У поздравној речи др Јуре Гашпарич, државни секретар Министарства образовања, науке и спорта, осврнуо се на значај

ИЗУМ-а и COBISS-а, у области библиотекарства чија делатност није само призната у Словенији, већ и широм региона, а то је препознао и УНЕСКО. COBISS и SICRIS су два врло важна дела у мозаику словеначке Стратегије развоја 2030 која се фокусира на квалитет живота свих, јер је едукативна стратегија која повезује науку и економију, од кључног значаја за пренос знања. Осврнуо се и на то да је COBISS Meet 2020 вероватно највећи догађај у области науке које је ИЗУМ организовао протеклих година заједно са словеначким универзитетима. Академска дигитална колекција Словеније представља релевантан национални информативни ресурс за истраживаче. То ће бити јединствена платформа за приступ пуном тексту чланака у електронским информативним изворима на које су универзитети и њихове библиотеке претплаћени појединачно, што ће у значајној мери побољшати целокупно искуство корисника. ИЗУМ припрема заједнички национални репозиторијум отворене науке назван dCOBISS. Он омогућава чување публикација истраживача које су у отвореном приступу, као и податке о пројектима. Надоградња националне инфраструктуре – НРС RIVR – суперкомпјутер ће почети са радом на пролеће 2021, а њиме ће бити ојачани национални рачунарска капацитети високих перформанси који су потребни за иновације како у истраживачком, тако и у подручју економије.

Затим се обратио Бахану Неупане, представник УНЕСКО-а, који је у свом излагању направио осврт на сарадњу УНЕСКО-а и ИЗУМ-а. Извршни одбор УНЕСКО-а почео је да разматра рад ИЗУМ-а пре осам година. Споразум између владе Словеније и УНЕСКО-а потписан је у септембру 2012. на шест година, међутим Центар је постао оперативан тек 14. јануара 2020. године под покровитељством УНЕСКО-а. Овај напор ће имати директан утицај на размену знања између различитих земаља. Затим је додао да ИЗУМ треба настави да ради на кључном питању утицаја на политике стварања, процесирања и архивирања знања, те се мора уметнути између места где знање постоји и оних на којима не постоји. Он је нагласио да ИЗУМ треба да шири знање кроз радионице, разне начине претресања идеја и и друге процесе размене знања. Ствари које је учинио ИЗУМ могу да послуже као пример добре праксе за земље у региону, а и шире. Посебан нагласак и похвалу је ставио на очување вишејезичности у оквиру система COBISS.net.

У излагању Мирана Петека имали смо прилику да слушамо о дигиталном репозиторијуму COBISS, или краће и dCOBISS, који је нови производ ИЗУМ-а, који је инсталиран и доступан свим библиотекама

у систему COBISS-а. Ово је нова веб апликација ИЗУМ-а намењена чувању дигиталног садржаја који се користи у библиотекама. Затим је објаснио везу са словеначким делом репозиторијума који је везан за словеначке библиотеке. Пријава на апликацију је аутономна и тесно интегрисана са COBISS3/Каталогизацијом, COBISS+ и Sicris. Све библиотеке у COBISS.net ће имати приступ централном спремишту, с тим да библиотекари из различитих система имају приступ само претраживачком делу спремишта, а библиотекари из локалног система државе чланице могу да похрањују записе и датотеке у своје централно спремиште и да их претражују. У процесу рада, библиотекар најпре креира библиографски запис у сегменту COBISS3/Каталогизација, а библиографском запису се у dCOBISS-у додају подаци о ауторским правима и датотеке дигиталних садржаја које су у отвореном приступу. Словеначке истраживачке агенције као и друге истраживачке академске институције подржавају припрему за аналитику отвореног приступа локалног спремишта која тек треба да буде функционална. Намера је да подаци из локалног спремишта буду унети у академско спремиште које ће бити део dCOBISS-а у отвореном приступу. Такође, све библиотеке ће имати приступ тестном окружењу dCOBISS, где ће библиотекари са корисничким именом и лозинком моћи да вежбају, тестирају и уче, а и где ће моћи да испробају нове надограђене функције.

Јанита Тацер Слана је представила новости у COBISS+, а то су наручивање чланака похрањених у COBISS+ у последње две године. Развој, надоградња и ажурирање COBISS-а новим функцијама су константне, а узимају у обзир примедбе библиотекара и корисника. Представљене су нове могућности у Мој профил COBISS, које се тичу израде списка резултата претраживања и ограничавања аспеката претраживања. Било је речи о последњој највећој, надоградњи која се односи на адаптацију за слепе и слабовиде, а такође је говорила и о корисничком интерфејсу за истраживаче. Сви портали раде на заједничкој софтверској платформи COBISS+. Предност за корисника се огледа у томе да када унесу термин за претрагу, претражују и COBISS и е-ресурсе у пуном тексту. Да би се побољшале корисничке услуге укључена је опција извоза у Excel-у. Јанита Тацер Слана осврнула се и на то како доћи до различитих група корисника брже, лакше и свеобухватније.

У излагању Тадеје Брешар, руководиоца библиографске контроле у ИЗУМ-у, имали смо прилику да се упознамо са општом листом предметних одредница COBISS. Назначила је да у заједничком каталогу

постоји 5,5 милиона записа, а отприлике око 83% од ових записа садржи бар једну предметну одредницу. Половина записа је креирана у Националној и унверзитетској књижници (NUK) и другим словеначким библиотекама, а само 6% записа садржи предметну одредницу из неке друге листе предметних одредница (Library of Congress Subject Heading, REMEAU и сл.). Листа предметних одредница створена је у облику нормативне базе података која омогућава јединствен систем именовања предметних одредница за цео систем COBISS. То значи да би све библиотеке могле да користе исти списак ових предметних одредница. База података је обезбеђује контролисани унос у пољу предметне одреднице, што значи да се више не могу унетосити погрешне предметне одреднице, нити их различити каталогизатори могу обликовати на различите начине. Контрола веродостојности омогућава контролу синонима, што значи да су сви нормативни записи опремљени нормативним и варијантним облицима, а сви каталози при додавању предметних одредница могу и морају да користе само нормативни облик предметне одреднице. Структурирана је општа листа предметних одредница, а записи у њима су повезани хијерархијски, како са надређеним, тако и са подређеним терминима.

Други дан конференције започела је Тања Туршек која ради у Одељењу за развој локалних апликација COBISS3. У свом излагању представила је неке нове функције и ажурирања које су уграђене у COBISS3. Представила је пар функција најактивнијег софтверског модула Позајмица, и то исправке које се односе на бржи приступ најчешће коришћеној грађи, претрагу чланова по кључу или имену, приступ фонду уџбеника, позајмљивање е-књига, резервацију путем COBISS+ или mCOBISS.

Тања Журан Путора је у свом излагању представила нову генерацију ревизије која је имплементирана 2019. године и названа је онлајн инвентар, а изводи се у производном окружењу COBISS3. Развој се одвијао у два корака – прво су припремали решење за ревизију монографија, па је додата могућност за ревизију серијских публикација. Апликација омогућава проналажење материјала на полици на основу инвентарних бројева, а преко бар-код читача.

У развоју апликације mCOBISS Боштјан Батич има улогу програмера, осим тога ради на промоцији и обуци корисника за употребу мобилне апликације. Његова презентација односила се на јединствену апликацију mCOBISS у оквиру мреже COBISS.net. Од 2013. године у ИЗУМ-у су започели развој мобилне апликације mCOBISS, пројекат

је суфинансирала Европска Унија под покровитељством Министарства за образовање, науку и спорт Словеније. Циљ за подношење пројекта био је подстицање развоја ширег избор е-услуга и мобилних апликација у различитим облицима информационог друштва, као и успостављање нових технолошких концепата који ће допринети транзицији у модерно и ефикасно друштво знања.

У излагању Сергеја Лаха имали смо прилику да се упознамо са COBISS-ом као друштвеном мрежом. Он у ИЗУМ-у ради на пословима односа са јавношћу, на обуци корисника и изградњи односа са корисницима, а члан је уредништва неколико ИЗУМ-ових друштвених медија. Лах је истакао да је 2003. године потписан Споразум о мрежи COBISS.net и размени библиографских записа између националних библиотечких информационих система у земљама Западног Балкана. COBISS.net развија информациони систем библиотека, повезује осам држава и укупно 1.400 библиотека, а преко 5.800 библиотекара је повезано и сарађује. У мрежи се користе информационе технологије на осам националних језика и два писма и то све представља огромну мрежу библиотекара и корисника библиотека који се могу повезати и на друштвеним мрежама које води и одржава ИЗУМ.

Трећег дана Миран Петек је у свом излагању, представио пројекат ADZ који се односи на Академску дигиталну збирку Словеније. Реч је о пројекту који је почео да се реализује током прошле године кад је инсталирана прва верзија тзв. *discovery* алата на платформи COBISS+. Ова платформа први пут је представљена на конференцији COBISS 2014. године, затим је 2016. године инсталиран COBISS+, дотадашњи COBISS/ОПАС замењен је платформом COBISS+. Тако је постављена основна платформа која користи функционалност алата *discovery* за истраживачко и академско окружење. Намера је да овај портал, постане портал Академске дигиталне збирке Словеније, а планиран је развој посебног дела COBISS+, који ће бити намењен истраживачкој сфери.

У излагању Романе Мухвич Шумандл, која је вођа Сектора управљања сервисима, имали смо прилику да се упознамо са новим функцијама корисничке подршке. Пројекат укључења школских библиотека у COBISS систем је завршен 2018. и представљен је у Савезу библиотекарских друштва. За потребе школских библиотека су извели поједине конверзије, јер су школске библиотеке радиле у другим системима, као што су WinKnj или Školska knjižnica. Затим је представљена надоградња програмске опреме опцијом Грађа у карантину, која је у спрези са радом библиотека у ситуацији

проузрокованој пандемијом вируса Ковид-19. Напоменула је да се повећао упис корисника путем интернета услед пандемије. Додала је и да се са комерцијалне базе података Oracle прешло на релационалну базу отвореног кода PostgreSQL. Занимљив податак је да се у COBISS.net за две године укључило укупно 95 библиотеке.

Бојан Шток, вођа одељења за развој софтвера, нас је упознао са архитектуром система COBISS и преласком на платформу отвореног кода. Осврнуо се на архитектуру COBISS система каква је данас и како се технологија мењала у развоју COBISS2 и COBISS3, као и каква нас технологија очекује у развоју софтвера COBISS4. Уместо виртуелних сервера користиће технологију контејнера. До технолошких промена ће долазити у зависности од развоја технологије у будућности. Тренутно је планирана употреба технологије JakartaEE, MicroProfile, који су наследници JAVE EE и омугућавају бржи развој апликације.

Гордана Мазих у свом излагању је говорила о идејној основи и прелиминарним резултатима анализе обављене у оквиру задатка креирања нормативних записа за дела у оквиру система COBISS. Она је истакла потребу да се код креирања нормативног записа за одређени наслов тј. дело направе релације са свим издањима и преводима дела на различите језике. Нормативни запис за дело повезан је и са нормативним записима за ауторе и са библиографским записима за његова различита издања и преводе. Нормативни записи за дела су попут слагалице која гради ширу слику о делима. Вођење оваквих нормативних записа ставља интелектуални и уметнички садржај на највиши ниво апстракције.

Сваког дана одржавала су се предавања, која су укупно трајала два сата, праћена дискусијом. Да би аудиторијум што боље разумео предаваче, сва предавања су била преведена на три језика. Модератори су били Боштјан Батич и Бранко Курњак.

Овај виртуелни скуп је био одлична прилика за расправу о идејама и заједничким плановима за будућност. Осим праћења стручног рада, размене мишљења, примера добре праксе и достигнућа са предавачима, колегама и другим учесницима, овај догађај је уједно био одлична прилика за друштвено умрежавање, што нам показује више од 1.500 коментара, поздрава, питања и одговора који су током трајања догађаја низали на страници Конференције. Презентација нових приступа и новости, које је ИЗУМ увео у последњем периоду, наишле су на добар одзив, што се исказало у дискусијама са предавачима. Заинтересованост за учешће у овом догађају је била велика. Поздравна реч и сва

предавања, укључујући дискусије, могу се погледати на You Tube каналу COBISS – poti do znanja, IZUM Maribor.¹

1. Видео материјали конференције

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 прештампавати 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 200 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 L^AT_EX, 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 L^AT_EX ће обавити редакција часописа.
2. Радови писани на српском језику треба да буду откуцани ТИРИЛИЧНИМ 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 100 карактера. За све наслове, 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 утјежђености.
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>).