УНИВЕРЗИТЕТ У БЕОГРАДУ
ФИЛОЛОШКИ ФАКУЛТЕТ

Датум 18. I 2016.

ИЗВЕШТАЈ О ПОДОБНОСТИ ТЕМЕ, КАНДИДАТА И МЕНТОРА ЗА ИЗРАДУ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

I ПОДАЦИ О КОМИСИЈИ

1. Датум и орган који је именовао комисију

Одлуком 2805/1 од 2. XII 2015. Научно-наставног већа Филолошког факултета Универзитета у Београду образована је комисија за оцену подобности докторанда Јелене Андоновски „Мрежа отворених података и језички ресурси у процесу изградње српско-немачког литерарног корпуса".

2. Састав комисије са назнаком имена и презимена сваког члана, звања, назив уже научне области за коју је изабран у звање, датум избора у звање и назива факултета, установе у којој је члан комисије запослен:

1. др Цветана Крстев, редовни професор, Филолошки факултет Универзитета у Београду, библиотечка информатика, 20. V 2014.
2. др Милош Утвић, доцент, Филолошки факултет Универзитета у Београду, библиотечка информатика, 18. XI 2014.
3. др Јелена Костић-Томовић, ванредни професор, Филолошки факултет Универзитета у Београду,
4. германистика, немачки језик, 3. IX 2013.
   др Душко Витас, ванредни професор, Математички факултет Универзитета у Београду, рачунарство и информатика, 8. VII 2011.
5. др Ранка Станковић, ванредни професор, Рударско-геолошки факултет Универзитета у Београду, математика и информатика, 19. V 2015.

II БИОГРАФИЈА КАНДИДАТА

Јелена Андоновски је рођена 21.09.1987. у Лесковцу. Основне студије је уписала на Филолошком факултету Универзитета у Београду школске 2006/2007. године, где је на Катедри за библиотекарство и информатику дипломирала 2010. године са просечном оценом 9,50. Школске 2010/2011. године уписала је мастер студије на истој Катедри и новембра 2011. године одбранила мастер рад „Европеана – врата до европског културног наслеђа". Мастер студије завршила је са просечном оценом 10,00. Школске 2012/2013. уписала је докторске студије Језик, књижевност, култура на Филолошком факултету Универзитета у Београду, модул Култура.

Јелена Андоновски је од новембра 2011. године запослена у Универзитетској библиотеци „Светозар Марковић" где је од децембра 2010. радила као волонтер. Радила је прво у Одељењу за развој и одржавање библиотечког система, дигитализацију фондова и културну делатност на пословима дигитализације фондова и креирања метаподатака, као и развијања и примене нових технологија у области библиотекарства. Од јуна 2012. године ради у Одељењу за обраду библиотечког материјала на пословима каталогизације и предметне и стручне класификације монографских публикација (стране и домаће књиге). Стручни испит у библиотечко-информационој делатности положила је новембра 2011. године. Од почетка рада у Библиотеци прошла је курсеве: COBISS3/фонд, COBISS3/Каталогизација, COBISS3/Каталогизација-монографске публикације.

Радила је на успостављању дигиталног репозиторијума Therefore (комплетно пословно решење за унапређење пословања библиотека у области рада са дигиталним документима), на креирању шеме за метаподатке за овај репозиторијум који је обухватао и послове мапирања метаподатака из COMARC формата. Учествовала је у организацији и одржавању курсева, предавања, презентација и радионица у области дигитализације и коришћења електронских извора информација.

Од фебруара 2012. године координира рад Аустријске библиотеке при Универзитетској библиотеци, која је поклон владе Републике Аустрије. Ради на пословима каталогизације и предметне и стручне класификације

публикација које стижу за фонд ове библиотеке и остварује сарадњу са колегама из Аустрије. У оквиру ове сарадње, преко Österreichische Gesellschaft für Literatur, била је као стипендиста на студијском бораваку у Бечу од 08-14.06.2014. у току кога је присуствовала радионицама везаним за рад Аустријских библиотека.

Јелена Андоновски је учествовала у два европска пројекта из области дигитализације, *Europeana Libraries (2011-2012)* и *Europeana Newspapers (2012-2015)*, на пословима дигитализације грађе и креирања метаподатака.

Јелена Андоновски је администратор електронског издања часописа *Инфотека: часопис за дигиталну хуманистику* од првог броја за 2014. годину. Уређује електронско издање часописа, ради на одржавању веб сајта часописа, успоставила је и тестира рад система за електронску пријаву радова (Open Journal System - OJS), остварује комуникацију са ауторима и координира техничку припрему бројева.

У току рада у Библиотеци Јелена Андоновски је учествовала на бројним националним и домаћим стручним и научним конференцијама (у организацији Филолошког факултета, Националног центра за дигитализацију, Библиотекарског друштва Србије, итд.). Објавила је научне и стручне радове у зборницима са домаћих и међународних конференција. Један је од организатора акредитованог курса „Метаподаци у савременом библиотекарству“ који је одржан у Универзитетској библиотеци 2015. години.

Према споразуму о сарадњи који је Универзитетска библиотека потписала са Гете институтом у Београду Јелена Андоновски је била стипендиста Гете института за учење немачког језика. У оквиру исте сарадње је као стипендиста била у Дизелдорфу четири недеље на усавршавању немачког језика. Координира сарадњу између Универзитетске библиотеке “Светозар Марковић” и Гете института.

Јелена Андоновски је члан и један од оснивача Друштва за језичке ресурсе и технологије Универзитета у Београду (JePTex - http://jerteh.rs/) која делује на пољу развоја електронских речника српског језика, српског језика на интернету, обогаћивања дигиталног корпуса српског језика, видљивости српског језика у мрежном окружењу.

III БИБЛИОГРАФИЈА КАНДИДАТА

1. Nedeljkov, Gordana, Jelena Andonovski and Zoran Bajin, „Digitization and interculturalism – University library “Svetozar Marković” in the project “Itinéraires balkaniques. Collection numérique en réseau”“, u *Pregled NCD* 24 (2014): 75-79, http://elib.mi.sanu.ac.rs/files/journals/ncd/24/ncd24075.pdf
2. Dakić, Nataša and Jelena Andonovski, „Developement of a new format EDM for metadata ingestion in Europeana“, u *Pregled NCD* 23 (2013): 11-21,http://elib.mi.sanu.ac.rs/files/journals/ncd/23/ncd23011.pdf
3. Андоновски, Јелена, Биљана Ђорђевић и Марија Лукић, „Мултимедијални документ – Александар Аца Поповић”, у *Књига и језик у развоју савременог друштва, Београд, 24-26. септембар 2010*, ур. Александра Вранеш, Љиљана Марковић, стр. 495-500. Београд: Филолошки факултет Универзитета, 2011
4. Дакић, Наташа и Јелена Андоновски, „Интегрисање метаподатака у Европеану коришћењем формата ESE“, у *Хоризонти светског и европског библиотекарства у дигиталном добу: зборник радова са међународне научне конференције, Београд, 27-28. октобар 2011*, ур. Весна Црногорац, Весна Ињац, стр. 67-80. Београд: Библиотекарско друштво Србије, 2012.
5. Андоновски, Јелена и Гордана Недељков, „Метаподаци у пројекту „Библиотеке Европеане“ – искуство двогодишњег пројекта“, у Културе у дијалогу. Културна дипломатија. Књ. 3, ур. Александра Вранеш, Љиљана Марковић, стр. 275-287. Београд: Филолошки факултет Универзитета, 2013.
6. Андоновски, Јелена и Александра Тртовац, „Улога дигиталног корпуса из библиотекарства и информатике у развоју речника српског језика“, у *Србија између истока и запада : наука, образовање, култура, уметност : тематски зборник у 4 књиге. Књ. 4, Језици Балкана у компаративном и интердисциплинарном контекст*, ур. Александра Вранеш, Љиљана Марковић, стр. 227-242. Београд: Филолошки факултет Универзитета, 2014.
7. Filipi-Matutinovic Stela; Andonovski, Jelena: „Europeana Newspapers project contribution to the freedom of information: finding out about Nikola Tesla from historical newspapers“. *II INFORUM 2013: 19th Annual Conference on Professional Information Resources, Prague 21-22 May 2005* [online]. Praha: Albertina icome Praha, 2013. Dostupno na:http://www.inforum.cz/archiv/inforum2013/en/proceedings/. ISSN 1801-2221
8. Trtovac, Aleksandra and Jelena Andonovski, “Enrichment of Morphological Dictionary of MWUs from Library and Information Science”, u *Natural Language Processing for Serbian: Resources and Applications*, ed. Gordana Pavlović-Lažetić, Cvetana Krstev, Ivan Obradović, Duško Vitas, pp. 27-40. Belgrade: Faculty of Mathematics, 2014.
9. Sofronijevic, Adam, Natasa Dakic and Jelena Andonovski, „Metadata in Serbia: describing library collections, towards a more Semantic web“, u *Knowledge discovery, transfer and managemant in Information Age*, ed. Murray E. Jennex (San Diego: IGI Global Publishing, 2014)

IV ОЦЕНА ДА ЈЕ КАНДИДАТ ПОДОБАН ДА РАДИ ДИСЕРТАЦИЈУ

Увидом у биографију и библиографију кандидата, комисија сматра да кандидаткиња Јелена Андоновски испуњава све услове за израду докторске дисертације.

V ОЦЕНА ПОДОБНОСТИ ПРЕДЛОЖЕНОГ МЕНТОРА

НАПОМЕНА: Кликните да бисте почели унос текста.

За ментора је предложена др Цветана Крстев, редовни професор
Филолошког факултета у Београду, Катедра за Библиотекарство и информатику.

Библиографија ментора са најмање 5 јединица релевантних за област из које се ради докторска дисертација.

1. Cvetana Krstev, Duško Vitas, "Konkordancije paralelizovanih tekstova", u *Zbornik radova XXXVIII konferencije ETRAN, Niš, juni 1994.*, ed. Slobodan Lazović, pp. 229-230, Društvo za elektroniku, telekomunikacije, računarstvo, automatiku i nuklearnu tehniku, Beograd, 1994.
2. Duško Vitas, Cvetana Krstev, "A lexical approach to text alignment using Intex", in *Intex pour la linguistique et le traitement automatique des langues*, eds. Claude Muller, Jean Royaute, Max Silberztein, pp. 249-263, Presses Universitaires de Franche Comté, Besancon, 2004.
3. Duško Vitas, Cvetana Krstev, "Structural derivation and meaning extraction: a comparative study on French-Serbo-Croatian parallel texts", in *Meaningful Texts: The Extraction of Semantic Information from Monolingual and Multilingual Corpora*, eds. Geoff Barnbrook, Pernilla Danielsson, Michaela Mahlberg, pp. 166-178, The University of Birmingham Press, Birmingham, 2005.
4. Duško Vitas, Cvetana Krstev, "Literature and Aligned Texts", in *Readings in Multilinguality*, eds. Milena Slavcheva, Galia Angelova and Kiril Simov, pp. 148-155, Institute for Parallel Processing, Bulgarian Academy of Sciences, Sofia, Bulgaria, 2006.
5. Duško Vitas, Cvetana Krstev, Eric Laporte, "Preparation and exploitation of Bilingual Texts", in *Lux Coreana*, No. 1, pp. 110-132, Han-Seine, 2006.
6. Cvetana Krstev, *Processing of Serbian – Automata, Texts and Electronic dictionaries*, Faculty of Philology, University of Belgrade, Belgrade, 2008.
7. Duško Vitas, Cvetana Krstev, "O paralelnim korpusima, a posebno o beogradskim paralelnim korpusima i načinu njihove eksploatacije", Sonderdruck aus: *Die Unterschide zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen (Razlike izmedju bosanskog/bošnjačkog, hrvatskog i srpskog jezika)*, ed. Branko Tošović, pp. 630-649, Münster, Germany, LITVerlag, 2008.
8. Cvetana Krstev, Duško Vitas, "An Aligned English-Serbian Corpus", In: *ELLSIIR Proceedings (English Language and Literature Studies: Image, Identity, Reality), Volume I, Belgrade, 4-6 December 2009*, eds. N. Tomović & J. Vujić, pp. 495-508, Faculty of Philology, University of Belgrade, ISBN 978-86-6153-005-0, 2011.
9. Duško Vitas, Ljubomir Popović, Cvetana Krstev, Ivan Obradović, Gordana Pavlović-Lažetić and Mladen Stanojević. *Српски језик у дигиталном добу -- The Serbian Language in the Digital Age*, Springer, in "META-NET White Paper Series", Georg Rehm and Hans Uszkoreit (Series Editors), 2012. ISBN (13) 978-3-642-30754-6.
10. Duško Vitas and Cvetana Krstev. "Processing of Corpora of Serbian Using Electronic Dictionaries". in *Prace Filologiczne*, vol. LXIII, pp. 279-292, Warszawa, 2012. ISSN 0138-0567.
11. Duško Vitas and Cvetana Krstev. "Construction and Exploitation of X-Serbian Bitexts". In Cristina Vertan and Walther v.Hahn (eds.) *Multilingual Processing in Eastern and Southern EU Languages: Low-Resourced Technologies and Translation*, pp. 207-227, Cambridge Scholars Publishing, 2012. ISBN (13) 978-1-4438-3878-8.

Комисија сматра да проф. Др Цветана Крстев **испуњава све услове за ментора ове докторске дисертације за област** библиотечка информатика и рачунарска лингвистика.

VI ОЦЕНА ПОДОБНОСТИ ТЕМЕ:

ОЦЕНА:

1. **Формулација назива тезе (наслова)**

Комисија закључује да је радни назив тезе „Мрежа отворених података и језички ресурси у процесу изградње српско-немачког литерарног корпуса" прикладан и да добро репрезентује суштину предложене теме истраживања.

2. Предмет (проблем) истраживања

Опис теме: Предмет истраживања ове докторске дисертације су паралелни корпуси и њихово коришћење у различитим областима истраживања и проучавања језика као што су двојезична или вишејезична лексикографија, учење страних језика, процеси превођења и машинског превођења, истраживање терминологије, лингвистичка истраживања, упоредна изучавања два или више језика итд. У оквиру истраживања кандидаткиња ће представити начине израде и анализу карактеристика, предности и значаја паралелних корпуса, на примеру српско-немачког корпуса.

У самој дисертацији кандидаткиња ће се бавити поступцима израде и успостављања једног паралелног корпуса, методама његове припреме, обраде и упаривања (паралелизације), процесима претраживања и проналажења информација у оквиру самог корпуса, као и поступцима повезивања са другим изворима на вебу. Да би се објаснили поступци повезивања са другим изворима на вебу, предмет докторске дисертације обухватиће и описивање и испитивање принципа и значаја семантичког веба и технологија које се на њему базирају, као што су повезани (отворени) подаци (Linked (Open) Data) и системи за организацију знања (контролисани речници, тезауруси, фасетне класификације, онтологије). Кандидаткиња ће расветлити њихову улогу и значај у процесу семантичког повезивања података на вебу и илустровати их на текстовима на српском и немачком језику.

Паралелни корпуси представљају посебне врсте корпуса који се састоје од паралелних текстова (битекстова). Чине их један или више оригиналних текстова и њихови преводи на један или више језика. Најчешће се формирају од два текста исте садржине, углавном на два различита језика, од којих је један текст оригинал, а други представља његов превод. Поред тога постоје паралелни корпуси са текстовима на више језика, али и корпуси са паралелним текстовима на истом језику (подразумева се да се корпусни текст састоји из различитих превода истог дела на један језик). Кандидаткиња ће посебно представити оне постојеће паралелне корпусе који као један језик обухватају српски и указати на непостојање паралелног корпуса већег обима и високог квалитета који би обухватао и српски и немачки језик.

Кандидаткиња ће даље објаснити у чему се састоји процес упаривања тј. паралелизације на припремљеном корпусу текстова, те ће бити представљене неке програмске платформе које се користе за поступке паралелизације и генерисања крајњих резултата. Како би процес паралелизације уопште био започет, текстови морају бити у електронском облику. Сва дела припремљена за потребе изградње корпуса који је предмет предложене докторске дисертације налазе се у аналогном облику, те ће прва фаза рада обухватити процесе сканирања текстова и њихово претварање у електронски облик. Коначни електронски облици текстова који ће бити укључени у овај корпус биће кориговани, припремљени у складу са важећим међународним и професионалним стандардима и снабдевени одговарајућим метаподацима, то јест структурираним информацијама које описују, објашњавају, лоцирају или на други начин чине лакшим проналажење, коришћење или управљање изворима информација. Метаподаци који ће бити придружени делима уграђеним у паралелни корпус биће, по својој природи описни (библиографски), структурални и административни. Примера ради, неки од административних података односе се на ауторска права, то јест регулишу права приступа и коришћења како метаподатака тако и самих дигиталних објеката.

Посебан део истраживања биће посвећен семантичком повезивању библиографских и структурираних метаподатака из анализираног корпуса са релевантним изворима на вебу. У овом делу дисертације најпре ће бити објашњен темин „семантички веб“ и његове карактеристике, формат који је развијен за опис семантичке повезаности података – RDF (Resource Description Framework), технологије и стандарди на којима се заснива - Linked (Open) Data, као и терминолошки ресурси који су део семантичког веба и омогућавају његову реализацију. Кандидаткиња ће такође анализирати одабране иницијативе и пројекте, углавном међународног карактера од којих неки укључују и домаће истраживаче и институције, који се заснивају на принципима семантичког веба, представиће њихове моделе података, начине претраге и приступа информацијама, као и предности овакве структуре за различите заједнице. Затим ће уследити представљање ресурса који ће се користити за повезивање података, а који ће обухватити библиографско-каталошке базе података и нормативне базе података, како за српско тако и за немачко говорно подручје. Након тога биће описан, представљен и анализиран начин примене ове технологије на корпусу текстова који је предмет дисертације. Коначно, кандидаткиња ће представити и алате које ће користити у предложеном истраживању од којих су многи развијени за српски језик и потребе других апликација.

Комисија закључује да је предложени предмет истраживања својом сложеношћу, актуелношћу и значајем у областима дисертације.

3. Библиографија прелиминарних истраживања

Кандидаткиња Јована Андоновски је приложила библиографију од 30 библиографских јединица релевантних за истраживање. Овде наводимо најзначајније јединице:

Antoniou, Grigoris and Frank Van Harmelen. 2008. *A Semantic Web Primer*. USA: Massachusetts Institute of Technology.

Caimi, Federico. *Ontology and Instance Matching for the Linked Open Data: thesis*. Chicago: University of Illinois, 2012.

Erjavec, Tomaž. "A TEI encoding of aligned corpora as translation memories." In *Proceedings of the EACL Workshop on Linguistically Interpreted Corpora*, vol. 99, pp. 49-60. 1999.

Gale, William A., and Kenneth W. Church. "A program for aligning sentences in bilingual corpora." *Computational linguistics* vol. 19, no. 1 (1993): 75-102.

Intner, Sheila S., Susan S. Lazinger and Jean Weihs. "Metadata and its impact on libraries". Westport, Conn. : Libraries Unlimited, 2006.

Klementiev, Alexandre, and Dan Roth. "Weakly supervised named entity transliteration and discovery from multilingual comparable corpora." In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pp. 817-824. Stroudsburg: Association for Computational Linguistics, 2006.

Koehn, Philipp. "Europarl: A parallel corpus for statistical machine translation." In *MT summit*, vol. 5, pp. 79-86. 2005.

Kazai, Gabriela. 2011. „In Search of Quality in Crowdsourcing for Search Engine Evaluation." *Proceedings of the 33rd Proceedings of the European conference on Advances in Information retrieval*. 165-176.

Krstev, Cvetana i Duško Vitas. "An Aligned English-Serbian Corpus". In *ELLSIIR Proceedings (English Language and Literature Studies: Image, Identity, Reality), Volume I, Belgrade, 4-6 December 2009*, eds. N. Tomović & J. Vujić, pp. 495-508. Belgrade: Faculty of Philology, University of Belgrade, 2011.

Krstev, Cvetana, Ivan Obradović, Miloš Utvić and Duško Vitas, "A system for named entity recognition based on local grammars". *J Logic Computation* vol. 24, no. 2 (2014): pp. 473-489

Lawson, Ann. "Collecting, aligning and analysing parallel corpora." In *Small corpus studies and ELT: theory and practice*, pp. 279-309. Amsterdam/Filadélfia: John Benjamins, 2001.

McEnery, Anthony and Zhonghua Xiao. "Parallel and comparable corpora: what are they up to?" *In Incorporating Corpora: Translation and the Linguist. Translating Europe. Multilingual Matters*. Clevedon, 2007.

Moore, Robert C. "Learning translations of named-entity phrases from parallel corpora." In *Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics* vol. 1, pp. 259-266. Association for Computational Linguistics, 2003.

Mosavi Miangah, Tayebeh. "Applications of corpora in translation." *Translation Studies* 12 (2006): 43-56.

Obradović, Ivan, Stanković, Ranka and Utvić, Miloš. 2008. "Integrisano okruženje za pripremu paralelizovanog korpusa". In *Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen*, ed. B. Tošović, 563–578. Muenster: LitVerlag.

Paskaleva, Elena, and Stoyan Mihov. "Second language acquisition from aligned corpora." In *Proceedings of International Conference "Language Technology and Language Teaching*, pp. 43-52.

Teubert, Wolfgang. "Units of meaning, parallel corpora, and their implications for language teaching." *Language and Computers* vol. 52, no. 1 (2004): 171-189.

Vitas, Duško and Cvetana Krstev. "Literature and Aligned Texts". In *Readings in Multilinguality*, eds. Milena Slavcheva, Galia Angelova and Kiril Simov, pp. 148-155. Sofia: Institute for Parallel Processing, Bulgarian Academy of Sciences, 2006.

Yu, Liang. *Introduction to the Semantic Web and Semantic Web Services*. Boca Raton: Chapman & Hall/CRC, 2007.

4. Циљеви истраживања

Циљеви: Основни циљ истраживања кандидаткиње Јелене Андоновски је покретање израде паралелног српско-немачког корпуса, као дела Корпуса савременог српског језика, са посебним акцентом на повезивању информација и метаподатака са релевантним ресурсима на вебу. Ради остваривања основног циља дисертације, користиће се све описане методе и технологије у складу са савременим информатичким и информационим принципима и захтевима.

Хипотезе: Главна хипотеза докторске дисертације кандидаткиње Јелене Андоновски од којих ће полазити у истраживању је:

• паралелни немачко-српски корпус развијен коришћењем језичких ресурса и алата, снабдевен метаподацима и повезан применом технологија семантичког веба са релевантним системима и ресурсима на вебу у ширу мрежну структуру Linked (Open) Data представљаће лексички ресурс који се може користити у задацима који захтевају ресурсе високог степена доследности и употребљивости у најразличитијим областима обраде природних језика.

5. Очекивани резултати

Као главни резултат докторске дисертације кандидаткиње Јелене Андоновски очекује се израда ресурса – паралелног српско-немачког корпуса који ће бити доступан истраживачкој јавности за разноврсна истраживања из области језика и књижевности. Очекује се да ће повезивање изграђеног корпуса са другим ресурсима на мрежи омогућити како повећање видљивости дигиталних колекција које се претражују тако и лакши приступ разноврсним подацима који су релевантни корисницима паралелног ресурса. Такође, очекује се да се произведени корпус искористи као добар пример у библиотекарству с обзиром на значајну примену библиотечких ресурса у процесу повезивања података и понуди решење за повезивање садржаја и података у мрежи отворених података.

6. План рада

1. Опис структуре и иницијалног садржаја корпуса и поступак одабира литературе према унапред постављеним критеријумима. Корпус ће иницијално бити састављен од дела савремених српских писаца који су превођени на немачки језик и дела савремених немачких писаца који су превођени на српски језик. Критеријуми за одабир обухватају:
    a. *аутори*: писци који су писали углавном у другој половини 20. века;
    b. *врста дела*: прозни текстови, углавном романи, без поезије и драмских текстова;
    c. *доступност*: постојање издања дела на оба језика у библиотекама у Србији;
    d. *награђиваност аутора*: дела која су награђена неком од међународних или националних награда или дела која су прославила аутора или је било најчитанија;
    e. *обимност дела*: краћа дела и дела средње дужине.
2. Процес сканирања материјала и оптичког препознавања карактера (OCR) коришћењем доступних сканера и софтвера за OCR;
3. Анализа и пречишћавање текста коришћењем описаних језичких алата и технологија;
4. Поступак анотације припремљених текстова и њихово упаривање (паралелизација) коришћењем одговарајућег софтвера;
5. Процес припреме и производње метаподатака у одговарајућем формату;
6. Постављање корпуса на веб;
7. Повезивање метаподатака са релевантним изворима на вебу коришћењем технологија семантичког веба, Linked (Open) Data;
8. Вредновање добијених резултата.

7. Методе истраживања

Кандидаткиња Јелена Андоновски ће у свом раду користити осим дескриптивне методе која ће укључити описивање, прикупљање и сређивање података, и компаративну и аналитичку истраживачку методу која подразумева упоређивање, вредновање и интерпретацију добијених резултата и анализу података из истраживања других аутора.

VII ЗАКЉУЧАК СА ОБРАЗЛОЖЕНОМ ОЦЕНОМ О ПОДОБНОСТИ ТЕМЕ КАНДИДАТА: Кандидаткиња **Јелена Андоновски** је показала својим досадашњим научним радом да је оспособљена за писање докторске дисертације. Докторска дисертација кандидаткиње **Јелене Андоновски** на тему **„Мрежа отворених података и језички ресурси у процесу изградње српско-немачког литерарног корпуса"**, имајући у виду њен теоријски оквир и иновативност има велики значај: циљ истраживања кандидаткиње је да изучи с једне стране методе за израду паралелних корпуса, а с

друге стране методе за повезивање података на вебу, посебну у мрежу отворених повезаних података, и да ове методе примени у пракси за израду иницијалне верзије литерарног немачко-српског паралелног корпуса. Посебна вредност рада биће вредан ресурс који ће преко веба бити доступан широкој истраживачкој заједници. Предложени план рада, листа библиографских јединица релевантних за истраживање као и предложени план истраживања и очекивани резултати показују да кандидаткиња суверено влада методама које су потребне за овакву врсту научног истраживања.

НАПОМЕНА:

На основу наведених података о кандидату, Комисија констатује да је кандидаткиња **Јелена Андоновски** подобна за израду докторске дисертације.

На основу постављених циљева и очекиваних резултата, Комисија констатује да је тема **„Мрежа отворених података и језички ресурси у процесу изградње српско-немачког литерарног корпуса“** подобна за израду докторске дисертације у области библиотечке информатике.

На основу наведених података о предложеном ментору, Комисија констатује да је **др Цветана Крстев** подобна за ментора предложене докторске дисертације.

На основу изнетог, Комисија предлаже Наставно-научном већу Филолошког факултета Универзитета у Београду да донесе позитивне оцене о подобности кандидаткиње **Јелене Андоновски** и теме докторске дисертације **„Мрежа отворених података и језички ресурси у процесу изградње српско-немачког литерарног корпуса“**, а да се за ментора именује **др Цветана Крстев**, редовни професор Филолошког факултета у Београду.

ПОТПИСИ ЧЛАНОВА
КОМИСИЈЕ

1. др Цветана Крстев, ред. професор
2. др Милош Утвић, доцент
3. др Јелена Костић Томовић, ван. професор
4. др Душко Витас, ван. професор
5. др Ранка Станковић, ван. професор