

ИЗБОРНОМ ВЕЋУ ФИЛОЛОШКОГ ФАКУЛТЕТА УНИВЕРЗИТЕТА У БЕОГРАДУ

На основу Одлуке Изборног већа Филолошког факултета Универзитета у Београду бр. 1475/1 од 21.V 2014. године, а у складу са чланом 65 Закона о високом образовању Републике Србије и чланом 116 Статута Филолошког факултета Универзитета у Београду, изабрана је Комисија за припрему извештаја о кандидатима који су се јавили за избор доцента за ужу научну област Библиотека информатика, предмет Дигитални текст 2 (објављен у листу „Послови“ бр. 572 од 4. VI 2014. године). Комисија подноси Изборном већу Филолошког факултета Универзитета у Београду следећи

ИЗВЕШТАЈ

На конкурс за доцента за ужу научну област Библиотека информатика, предмет Дигитални текст 2, који је објављен у листу „Послови“ бр. 572 од 4. VI 2014. године, пријавио се један кандидат — др Милош УТВИЋ, асистент за ужу научну област Библиотека информатика на Филолошком факултету.

I. БИОГРАФИЈА И АНГАЖМАН У СТРУЦИ

Милош Утвић је рођен 29. јануара 1976. године у Прокупљу. Основну школу и гимназију је завршио у Куршумлији као ђак генерације. Дипломирао је 29. маја 2001. године на Математичком факултету у Београду на смеру рачунарство и информатика са просечном оценом 9,34 (девет и 34/100). Степен магистра рачунарства је стекао 18. фебруара 2008. године одбранивши магистарски рад са темом *Коначни аутомати у регуларној именској деривацији* на Математичком факултету Универзитета у Београду. Докторску дисертацију под називом *Изградња референтног корпуса савременог српског језика* је одбранио 28. априла 2014. године на Филолошком факултету Универзитета у Београду и стекао титулу доктора библиотечко-информатичких наука.

Од 3. децембра 2001. године ради као асистент-приправник на Катедри за библиотекарство и информатику на Филолошком факултету у Београду. У звање асистента на Филолошком факултету Универзитета у Београду је биран 15. октобра 2008. године и поново 12. октобра 2011. године. Сада је асистент за предмете *Информатика за библиотекаре 1–2, Информатички практикум 1–4, Дигитални текст 1–2, Базе података и библиотечки информациони системи, Проналажење информација, Мултимедијални документи* (основне студије) и *Електронско издаваштво и дигиталне библиотеке* (мастер студије).

Члан је Групе за језичке технологије Универзитета у Београду, као и Комисије за корпус Одбора за стандардизацију српског језика САНУ. Од програмских решења је развио десетак подсистема у оквиру Корпуса савременог српског језика, а у сарадњи са др Ранком Станковић и интегрисано окружење за припрему паралелизованих корпуса. Учествовао је на радионицама, конференцијама и симпозијумима и презентовао више радова.

Учествовао је на четири домаћа научна пројекта и на једном међународном пројекту. Био је члан комисије за одбрану великог броја мастер радова на Филолошком факултету Универзитета у Београду.

II. БИБЛИОГРАФИЈА

Научни и стручни рад кандидата др Милоша Утвића најбоље илуструје приложена библиографија, као и преглед других научних, стручних и наставничких активности.

1. Научни радови:

Академски радови:

Магистарски рад

1. Utvić, Miloš. 2008. Konačni automati u regularnoj imenskoj derivaciji. Magistarski rad, Matematički fakultet, Univerzitet u Beogradu. (105 str.)

Докторска дисертација

2. Utvić, Miloš. 2014. Izgradnja referentnog korpusa savremenog srpskog jezika. Doktorska disertacija, Filološki fakultet Univerziteta u Beogradu. (396 str.)

Радови у тематским зборницима међународног значаја (10)

3. Utvić, Miloš. 2006. Regular Derivation of Name Entities in Serbian. In *Proceedings of the 5th International Conference of Formal Approaches to South Slavic and Balkan Languages, FASSBL-5*, ed. S. Koeva and M. Dimitrova-Vulchanova, 257–264. Sofia: The Department of Computational Linguistics, Institute of Bulgarian Language. (264 p.)
4. 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. (ISBN: 978-3-8258-0144-1, 846 p.)
5. Krstev, Cvetana, Stanković, Ranka, Obradović, Ivan, Vitas, Duško and Utvić, Miloš. 2010. Automatic Construction of a Morphological Dictionary of Multi-Word Units. In *Advances in Natural Language Processing, 7th International Conference on NLP, IceTAL 2010, Reykjavik, Iceland, August 16-18, 2010*, ed. H. Loftsson, E. Rognvaldsson and S. Helgadóttir, 226–237. Vol. 6233. Berlin: Springer. (ISBN: 978-3-642-14769-2, 448 p.)
6. Utvić, Miloš. 2010. The Regular Derivation in Serbian: Principles and Classification Using NooJ. In *Applications of Finite-State Language Processing: Selected Papers from the 2008 International NooJ Conference*, ed. T. Váradi, J. Kuti and M. Silberstein, 142–151. Cambridge: Cambridge Scholars Publishing. (ISBN: 978-1-4438-2573-3, 1-4438-2573-5, 225 p.)
7. Utvić, Miloš, Obradović, Ivan, Krstev, Cvetana and Vitas, Duško. 2010. The Effects

of Multi-Word Tagging on Text Disambiguation. In *Proceedings of the 29th International Conference on Lexis and Grammar, September 15-18, 2010*, ed. D. Vitas and C. Krstev, 333–342. Belgrade: Faculty of Mathematics, University of Belgrade, Serbia.

(ISBN: 978-86-7589-080-5, 356 p.)

8. Krstev, Cvetana, Vitas, Duško, Obradović, Ivan and Utvić, Miloš. 2011. E-Dictionaries and Finite-State Automata for the Recognition of Named Entities. In *Proceedings of the 9th International Workshop on Finite State Methods and Natural Language Processing*, ed. A. Maletti and M. Constant, 48–56. Blois, France: Association for Computational Linguistics. (ISBN: 978-1-62276-464-8, 154 p.)
9. Yang, Ying, Utvić, Miloš, Pavlović-Lažetić, Gordana and Vitas, Duško. 2011. Named Entities in Chinese. In *Proceedings of the NooJ 2010 International Conference and Workshop, May 27-29, 2010, Komotini, Greece*, ed. Z. Gavrilidou, E. Chadjiapapa, L. Papadopoulou and M. Silberstein, 112–124. Democritus University of Thrace. (281 p. URL: http://synmorphose.compulaw.gr/joomlatools-files/docman-files/Zoe-Gav_BOOK_7.pdf)
10. Stanković, Ranka, Krstev, Cvetana, Obradović, Ivan, Trtovac, Aleksandra and Utvić, Miloš. 2012. A Tool for Enhanced Search of Multilingual Digital Libraries of E-journals. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation, LREC 2012, Istanbul, Turkey, 23–25 May 2012*, ed. N. Calzolari, K. Choukri, T. Declerck, M. U. Dogan, B. Maegaard, J. Mariani, J. Odijk and S. Piperidis, 1710–1717. (ISBN: 978-2-9517408-7-7, 4148 p. URL: http://www.lrec-conf.org/proceedings/lrec2012/pdf/375_Paper.pdf)
11. Stanković, Ranka, Utvić, Miloš, Vitas, Duško, Krstev, Cvetana and Obradović, Ivan. 2012. On The Compatibility Of Lexical Resources For NooJ. In *Automatic Processing of Various Levels of Linguistic Phenomena: Selected Papers from the 2011 International NooJ Conference*, ed. K. Vučković, B. Bekavac and M. Silberstein, 96–108. Cambridge: Cambridge Scholars Publishing. (ISBN: 978-1-4438-3711-8, 280 p.)
12. Zečević, Anđelka and Utvić, Miloš. 2012. An Authorship Attribution for Serbian. In *Local Proceedings of BCI'12, September 16–20, 2012, Novi Sad, Serbia*, ed. Z. Budimac, M. Ivanović and M. Radovanović, 109–112. Vol. 920. Novi Sad: Faculty of Sciences, University of Novi Sad. (ISBN: 978-86-7031-200-5, 150 p. URL: <http://perun.pmf.uns.ac.rs/bci2012/proceedings/local/p109-zecevic.pdf>)

Рад у врхунском међународном часопису (1)

13. Krstev, Cvetana, Obradović, Ivan, Utvić, Miloš and Vitas, Duško. 2014. A System for Named Entity Recognition Based on Local Grammars. *Journal of Logic and Computation* 24 (2): 473–489. (ISSN: 0955-792X. doi: 10.1093/logcom/exs079, URL: <http://logcom.oxfordjournals.org/content/24/2/473>)

Рад у научном часопису (1)

14. Утвић, Милош. 2011. Анотација Корпуса савременог српског језика. *Инфотека* XII (2): 39–51. (ISSN: 1450-9687. URL: http://infoteka.bg.ac.rs/PDF/Srp/2011-2/INFOTHECA_XII_2_Decembar_39-51.pdf)

Саопштење са међународног скупа штампано у изводу (1)

15. Utvić, Miloš. 2006. Derivation from Multi-word Proper Names Using NooJ. In *9th INTEX/NooJ Conference (Abstracts)*, ed. D. Vitas and M. Silberstein, 69–70. Belgrade, Serbia: Faculty of Mathematics, University of Belgrade, LASELDI, Université de Franche-Comté, Maison des Sciences de l'Homme Claude Ledoux. (ISBN: 86-7589-055-9, 84 p.)

Рад у часопису националног значаја (1)

16. Утвић, Милош. 2010. BibTeX - креирање персоналних библиографија. *Панчевачко читалиште: научни часопис за теорију и праксу библиотекарства* IX (17): 35–38. (ISSN: 1451-3048)

Рад у тематском зборнику националног значаја (1)

17. Утвић, Милош. 2008. Настава природно-математичке групе предмета у школској библиотеци. У *Школски библиотекар сарадник у настави*, ур. А. Вранеш, Љ. Марковић и П. Станојевић, 185–190. Београд: Филолошки факултет Универзитета у Београду, Библиотекарско друштво Србије. (ISBN: 978-86-86419-55-2, 208 стр.)

Остало:

18. Утвић, Милош. 2003. Регуларни изрази. *Свет речи* VII (15–16): 66–68. (ISSN: 1450-6211)
19. Утвић, Милош. 2007. Корпус савременог српског језика. *Свет речи* XI (23-24): 34-36. (ISSN: 1450-6211)

2. Пројекти:

Учешће на домаћим научним пројектима

1. 1743 *Интеракција текста и речника*, Министарство за науку Републике Србије, 2002-2004. године;
2. 148021 *Теоријско-методолошки оквир за модернизацију описа српског језика*, Министарство за науку Републике Србије и САНУ, 2005-2010. године;
3. III 178006 *Српски језик и његови ресурси: теорија, опис и примене*, Министарство за образовање и науку Републике Србије (у току од 2011. године);
4. III 47003 *Инфраструктура за електронски подржано учење у Србији*, Министарство за образовање и науку Републике Србије (у току од 2011. године).

Учешће на међународним научним пројектима

1. *Ресурси Средње и Југоисточне Европе* (енг. *Central and South-East European Resources*, skr. *CESAR*), 2011-2013. године (ICT Policy Support Programme, Grant agreement no.: 271022);

3. Софтверски алати и ресурси

Лексички ресурси

1. Корпус савременог српског језика доступан преко веба од 122 милиона текућих речи (коаутор);
2. Вишејезички корпуси (коаутор):
 1. *Паралелни српско-француски корпус*, чији је узорак доступан преко веба. Корпус се састоји од литерарних и новинских текстова у укупној дужини од 1,7 милиона текућих речи;
 2. *Паралелни српско-енглески корпус*, чији је узорак доступан преко веба. Корпус се састоји од литерарних и новинских текстова у укупној дужини од 4,5 милиона текућих речи;

Софтверски алати:

1. *Адаптација TNT-а за српски скуп етикета* (2004)
2. *ACIDE*, интегрисано окружење за припрему паралелизованог корпуса, (2007)
3. Развој више подсистема као подршка раду са Корпусом савременог српског језика и вишејезичним корпусима:
 - a. *Подсистем за вертикализацију текста и израчунавање статистике корпуса савременог српског језика*, (2003);
 - b. *Подсистем за администрирање неетикетираних и етикетираних корпуса (front-end за IMS CWB)*, (2003);
 - c. *Подсистем за проверу и корекцију текстуалних ресурса за корпус савременог српског језика*, (2003);
 - d. *Подсистем за администрирање паралелних корпуса (front-end за IMS CWB)*, (2004);
 - e. *Подсистем за конверзију текстуалних ресурса из формата MULTEXT-East у улазни формат IMS CWB*, (2004)
 - f. *Подсистем за генерисање фреквенцијара корпуса савременог српског језика сортираног а) алфабетски растуће, б) по суфиксима растуће, ц) по фреквенцијама опадајуће*, (2005)
 - g. *Подсистем за генерисање биграма и триграма корпуса савременог српског језика сортираних опадајуће по учестаности*, (2005)
 - h. *Подсистем за детекцију фразе (frozen expressions) у корпусу савременог српског језика*, (2005)
 - i. *Подсистем за одређивање расподеле лексема и типова савременог српског језика по дужини*, (2005)
 - j. *Подсистем за одређивање расподеле лексема по учестаности (тестирање Zipфовог закона)*, (2005)
 - k. *Подсистем за аутоматску сегментацију лексема*, (2005)
 - l. *geord*, аутоматски класификатор топонима на основу механизма регуларне деривације, (2006)
 - m. *Подсистем за екстракцију одабраних конкорданци из Корпуса савременог српског језика*, (2007)

- п. *Подсистем за приказивање информација о изворима текстова из Корпуса савременог српског језика*, (2007)

III. Преглед и мишљење о досадашњем научном и стручном раду кандидата

Научни рад

Истраживачки рад кандидата се одвија у оквиру Београдске групе за језичке технологије, која окупља истраживаче са више факултета Универзитета у Београду. Подручје језичких технологија је интердисциплинарно подручје у коме се информатичка истраживања пресецају са истраживањима језика. У овом оквиру, основне области научно-истраживачког рада кандидата др Милоша Утвића су првенствено везане за обраду природних језика, посебно за обраду српског, а затим и за проблем изградње језичких ресурса. Ова истраживања обухватају изградњу формализованих модела језика, њихову програмску имплементацију, а затим и изучавање њиховог понашања у информатизованој средини. Према увиду у приложени библиографију, највећи део објављених научних и стручних радова приказује резултате истраживања на подручју обраде природних језика и то из области изградње корпуса, рачунарске морфологије и морфосинтаксе, као и парцијалне синтаксичке анализе (именовани ентитети).

1. Корпуси. Кандидат је дао велики допринос изградњи више корпуса. Пре свега, ту је *Корпус савременог српског језика* који је доступан преко веба са сервера смештеног на Математичком факултету. Ово је електронски, синхрони и динамички корпус савременог српског језика, величине од 122 милиона речи. Корпус је анотиран на три нивоа: корпусним текстовима су придружене библиографске информације, део текстова је структурно анотиран, док су токени корпусних текстова аутоматски анотирани морфолошким информацијама о врсти речи и леми. Балансираност корпуса још увек није постигнута, али с обзиром да је корпус динамички и да се непрестано ажурира, овом циљу ће се тежити у будућности. Изграђено веб-сучеље које омогућује једноставну и напредну претрагу је у великој мери прилагођено потребама разноврсних корисника. Кандидат др Милош Утвић је наставио рад чланова Београдске групе за језичке технологије и деценију свог рада посветио изградњи и одржавању овог корпуса. У овом периоду кандидат је прикупљао и припремао хиљаде текстова за овај корпус с циљем да постигне одређен ниво репрезентативности. У истом периоду аутор је израдио велики број софтверских решења (набројаних у одељку II.3) којима се олакшава и унапређује, како прикупљање и анотација текстова за Корпус, тако и рад са самим Корпусом. Као резултат тога, Корпус савременог српског језика је данас признати и незаобилазни ресурс за рад многих лингвиста и филолога, како у Србији тако и у свету. Изградњи овог корпуса посвећена је и докторска дисертација кандидата (рад под бројем 2), а један посебан сегмент овог рада, анотација Корпуса, је описан у раду 14. Анотација је обављена на више нивоа: сваком тексту корпуса је, пре свега, придружена одговарајућа библиографска информација, а затим је, на основу морфолошког електронског речника српског језика и ручно анотираног корпуса од милион речи, прилагођен прогам TreeTagger и извршена морфосинтаксичка анотација целокупног Корпуса, тј. свим корпусним речима је придружена информација о врсти речи и леми. Преко ових информација је сада и омогућено претраживање Корпуса. Популарисању лингвистичких истраживања коришћењем корпуса, а посебно значају Корпуса српског језика за изучавања српског језика и књижевности, посвећен је рад 19.

Благи увод у регуларне изразе, једноставан механизам за формализацију једне класе формалних језика, дат је у раду 18. Умешно коришћење регуларних израза омогућава корисницима Корпуса да задају врло сложене упите.

Другу групу радова чине радови посвећени изградњи и експлоатацији паралелних корпуса. Кандидат је учествовао у изградњи више паралелних корпуса и колекција паралелизованих текстова поменутих у одељку II.3. Ради се о корпусима који садрже две или више варијанти истог текста, најчешће у виду изворног текста и његовог превода на други језик, али и вишеструких превода неког текста на исти језик. Припрема паралелизованих корпуса је још сложенија од припреме једнојезичних корпуса, те осим свих послова које је потребно урадити при припреми једнојезичних корпуса, треба извршити и паралелизацију на нивоу изабраног текстуалног сегмента. У раду под бројем 4 кандидат представља софтверски систем ACIDE који поједностављује процес контроле паралелизације, а паралелизовани текст представља у више излазних формата: TMX, TEI и HTML. Систем ACIDE представља надградњу постојећег програма за паралелизацију Xalign и нуди кориснику удобно графичко окружење. Кандидат је, такође, учествовао у пројекту изградње вишејезичне паралелизоване верзије романа Жила Верна „Пут око света за 80 дана” која данас укључује двадесетак језика, између осталих француски, енглески, српски, бугарски, грчки, албански, итд. Изузетност овог подухвата повећава обухваћеност кинеског језика. Овај подухват, који није био без изазова с обзиром на специфичност кинеског језика у односу на језике Европе (почев од кодирања па надаље), описан је у раду 9. Рад описује кораке у припреми паралелизованог текста, почев од припреме кинеске верзије текста, његове сегментације на реченице и паралелизације коришћењем система ACIDE. У раду под бројем 10 приказана је веб-апликација *Bibliša* која омогућава да се на вебу постављају проширени упити над паралелизованом верзијом чланака објављених у часопису *Инфотека*. За разлику од обичног упита, који тражи у колекцији текстова оне који дословно одговарају тексту упита, корисник система *Bibliša* може да изврши морфолошку и семантичку експанзију упита, као и његово превођење на други језик и да на тај начин обезбеди повољнију прецизност у претраживању.

2. Именовани ентитети. Задатак препознавања именованих ентитета је дефинисан 1997. као суштински истраживачки задатак на подручју екстракције информације који обухвата аутоматско препознавање властитих имена (посебно личних имена, назива организација и локација), као и препознавање датума, периода, процената и новчаних износа. Кандидат је сарађивао са екипом која је развила систем локалних граматика и каскада трансдуктора који користе морфолошке е-речнике српског језика, а служе да се за српски језик препознају широке класе именованих ентитета. Радови под бројем 8 и 13 разматрају из теоријског угла улогу локалних граматика и каскада коначних трансдуктора у препознавању именованих ентитета. Осим тога, у раду су приказани резултати евалуације изграђеног система на изабраној колекцији новинских текстова који показују да систем има високе перформансе како у погледу прецизности тако и у погледу одзива. Именованим ентитетима кандидат се бави и у раду 9 у коме се описује детекција именованих ентитета у кинеском, посебно временских израза, коришћењем како паралелизоване кинеско/енглеске верзије романа Жила Верна „Пут око света за 80 дана“ (описане у претходној тачки), тако и специјално изграђених речника и локалних граматика за кинески.

3. Полилексичке јединице. Аутоматска анализа текста се врши на нивоу простих речи (непрекинутих низова алфабетских симбола), док се, у формалном смислу, под полилексичким јединицама подразумевају непрекинути низови простих речи. Ова апроксимативна дефиниција омогућава да се прикупе и обраде различити

изрази и унесу у одговарајући тип морфолошког електронског речника са циљем да се отклоне вишезначности које се јављају током анализе која користи само речнике простих речи. Тако ће, на пример, одредница *црна овца*, која се у речнику простих речи јавља као низ од две засебне одреднице - *црн*, придев обележен као боја, и *овца*, именица која означава зоолошки појам, у речнику полилексичких јединица јавити као једна одредница и то као именица која означава особу. Слично, израз *из дана у дан* који се на нивоу простих речи анализира као низ *предлог именица предлог именица*, у анализи која укључује полилексичке јединице добија статус *прилога*. Променљиве полилексичке јединице постављају и проблем дефинисања њихових флективних својстава. Наиме, комбинација две именице која чини полилексичку одредницу може бити таква да се само прва или само друга именица мења или да се обе мењају. У наведеном примеру *црна овца*, уколико се не спецификује услов слагања између придева и именице, већ се допусти да се оне мењају независно, генераторни механизам производи на стотине некоректних облика. У светлу ових примедби, др Милош Утвић је у раду под бројем 5 описао принципе конструкције речника полилексичких јединица у српском, користећи програмски систем *Multiflex* који се темељи на механизму унификације. Кандидат је учествовао у изградњи посебног програмског подсистема који омогућава да се предвиде могуће флективне класе полилексичке јединице на основу њених компонената. У раду је дата и евалуација предложеног поступка. Сама метода је значајно убрзала рад на конструкцији речника полилексичких јединица. У раду под бројем 7 упоређени су резултати добијени у аутоматској морфолошкој анализи са и без укључивања речника полилексичких јединица. Приказан је степен смањења вишезначности на излазу из морфолошког анализатора у анализи са полилексичким јединицама у односу на анализу без њих.

4. Морфологија. Кандидат др Милош Утвић се у области рачунарске морфологије бавио посебно питањима деривације у српском језику чему је посвећена и магистарска теза кандидата (број 1). Кандидат је у овом раду предложио решења заснована на методама коначних аутомата за један аспект деривације, регуларну именску деривацију, која се уклапају у методологију анализе текстова на српском језику коришћењем електронских речника. На овим идејама кандидат је формулисао принципе за класификацију регуларне именске деривације увођењем појма *суперлеме*. Описивање суперлема коначним аутоматима омогућава систематско проширивање електронског речника српског језика чиме се побољшавају резултати анализе (рад број 6). Регуларна деривација у српском језику је посебно продуктивна у случају геополитичких властитих имена која су, пак, посебно важна за аутоматску анализу. Кандидат је у радовима 3 и 15 показао да регуларна деривација геополитичких имена добијена суфикасацијом над лемама које све припадају истој флективној класи (имају еквивалентно флективно понашање) генерише леме које припадају тачно одређеној флективној класи. Кандидат је посебну пажњу посветио случајевима регуларне деривације у којима се просте речи генеришу из полилексичких јединица. У раду под бројем 11 се анализирају недоследности у кодирању граматичких својстава у имплементацији електронских речника за више језика: енглески, француски, мађарски, пољски, хрватски, бугарски и српски, с посебним освртом на имплементацију тих речника под системом *NooJ*.

5. Остало. Кандидат се бавио једним посебним аспектом обраде природних језика у раду 12, а то је утврђивање ауторства текстова на српском језику. Проблем утврђивања ауторства се састоји у идентификовању аутора неког текста за који аутор није познат или нема сагласности око аутора (између више могућих). У решавању овог задатка користе се различите врсте знања о језику, пре свега, лексичко богатство, синтакса и семантика. Аутор у овом раду тестира применљивост познатих алгоритама

који се заснивају на употреби *n*-грама, као и модификован алгоритам заснован на слоговима, на српски језик и показује да оба ова приступа дају резултате који обећавају.

Део своје научне и стручне активности кандидат др Милош Утвић је посветио унапређењу наставе, посебно из информатичке групе предмета студентима библиотекаштва и информатике, и студентима Филолошког факултета уопште. У раду 16 промовише коришћење библиографског формата BibTeX, члана шире фамилије формата TeX и LaTeX, за формирање персоналних библиографија. Како школска библиотека и библиотекар могу помоћи ученицима у савладавању градива из групе природно-математичких предмета, коришћењем многобројних ресурса и он-лајн туторијала на вебу, кандидат излаже у раду 17.

Научна и стручна активност. На основу података из биографије овде истичемо да др Милош Утвић већ више од десет година развија и администрира Корпус савременог српског језика који у овом тренутку има више од 200 редовних корисника из целог света, а користи се и у настави на више факултета у земљи. Др Милош Утвић је члан Комисије за корпус Одбора за стандардизацију српског језика САНУ. Поред овога, др Милош Утвић у сарадњи са другим члановима Групе за језичке технологије Универзитета у Београду активно учествује у развоју алата и ресурса намењених обради и истраживању српског језика.

Наставна и наставничка активност. Др Милош Утвић је до сада држао вежбе из једанаест предмета на основним студијама и то *Информатика за библиотекарe 1 и 2, Информатички практикум 1-4, Дигитални текст 1 и 2, Проналажење информација, Базе података и библиотечки информациони системи и Мултимедијални документи*, а такође је укључен у наставу на мастер студијама. За потребе студената припремио је обиман наставни материјал који им је доступан преко веба и који се сваке године допуњује и ажурира. Предано је пратио израде великог броја обавезних семинарских радова и био члан комисије за одбрану више мастер радова. Одговорност према прихваћеним задацима испољава и у тимском раду, као сарадник пројекта *Српски језик и његови ресурси: теорија, опис и примене* обавља савесно разне послове везане за администрирање пројекта. Иако су наставне обавезе др Милоша Утвића изузетно обимне он се укључује и у друге облике и активности на Катедри и на Филолошком факултету, првенствено оне везане за акредитацију.

IV. Испуњеност других услова за избор

Увидом у приложену документацију, као и у рад и радове кандидата, Комисија констатује да су у потпуности испуњени формални (законски и статутарни) и суштински услови за избор др Милоша Утвића у звање доцента. Кандидат има један објављен рад у врхунском међународном часопису (на ISI листи), једанаест радова у тематским зборницима (10 у зборницима међународног значаја и један у зборнику националног значаја), два рада у домаћим научним часописима, један рад саопштен на скупу међународног значаја штампан у изводу и два стручна рада.

V. Предлог Комисије

На основу свега изложеног и што је Комисији познато, једногласно и свесрдно предлагемо Изборном већу Филолошког факултета Универзитета у Београду да

др Милоша Утвића изабере у звање доцента за ужу научну област Библиотека информатика, предмет Дигитални текст 2.

У Београду, 21. јула 2014.

Комисија:

(др Цветана Крстев, ред. проф.)

(др Александра Вранеш, ред. проф.)

(др Иван Обрадовић, ред. проф.)