

ИЗБОРНОМ ВЕЋУ ФИЛОЛОШКОГ ФАКУЛТЕТА УНИВЕРЗИТЕТА У БЕОГРАДУ

На основу Одлуке Изборног већа Филолошког факултета Универзитета у Београду бр. 2677/1 од 25.9.2013. 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 избор кандидата који су се јавили за избор редовног професора за ужу научну област Библиотека информатика, предмет Проналажење информација (објављен у листу „Послови“ бр. 538 од 9. X 2013. године). Комисија подноси Изборном већу Филолошког факултета Универзитета у Београду следећи

ИЗВЕШТАЈ

На конкурс за редовног професора за ужу научну област Библиотека информатика, предмет Проналажење информација, који је објављен у листу „Послови“ бр. 538 од 9. октобра 2013. године, пријавио се један кандидат — др Цветана КРСТЕВ, ванредни професор за ужу научну област Библиотека информатика на Филолошком факултету.

I. БИОГРАФИЈА И АНГАЖМАН У СТРУЦИ

Др Цветана Крстев рођена је 17. априла 1952. у Београду. Основну школу и гимназију завршила је у Београду. Дипломирала је на Природно-математичком факултету у Београду, Одсек за математику, смер за рачунарство и кибернетику, 1975. године. Магистрирала је на Математичком факултету Универзитета у Београду 1989. године одбравивши рад „Програмски системи за уређивање текста“ (ментор је био проф. др Недељко Парезановић). Докторирала је 1997. године на Математичком факултету Универзитета у Београду, а тема доктората је гласила „Један прилаз информатичком моделирању текста и алгоритми његове трансформације“ (ментор проф. др Гордана Павловић-Лажетић).

После кратке праксе као средњошколског наставника математике (1975), од 1976. до 1985. године ради прво као програмер, а затим као аналитичар у Рачунском центру Математичког института САНУ. Од 1985. до 1998. године је радила као стручни сарадник у Рачунарској лабораторији Математичког факултета Универзитета у Београду. 1998. године је изабрана у звање доцента на Катедри за библиотекарство и информатику Филолошког факултета Универзитета у Београду, а 2009. године је изабрана у звање ванредног професора. На основним студијама предаје предмете *Информатика за библиотекаре I и II*, *Информатички практикум I-IV*, *Дигитални текст I и II*, *Структуре информација I и II*, *Базе података и библиотечки информациони системи*, *Проналажење информација* и *Мултимедијални документи*. Ангажована је и на мастер и докторским студијама Филолошког факултета, на којима држи специјалне курсеве из области проналажења информација и рачунарске лингвистике.

Од 2007. године је главни и одговорни уредник часописа *Инфотека* (<http://www.zbus.rs/>) и члан је савета међународног часописа *Journal of the Bangladesh Association of Young Researchers*. Члан је Управног одбора и Извршног одбора COST-

акције IC1207 “PARSEME: PARSing and Multi-word Expressions. Towards linguistic precision and computational efficiency in natural language processing” који финансира Европска комисија у периоду 2013-2017.

Учествовала је на 13 домаћих научних и развојних пројеката и на 9 међународних и билатералних пројеката.

Члан је Комисије за корпус Одбора за стандардизацију српског језика. Обављала је функцију управник Катедре за библиотекарство и информатику Филолошког факултета. Члан је Савета Филолошког факултета Универзитета у Београду.

Била је члан програмског одбора на 20 међународних конференција и рецензент за часописе Slovenščina 2.0 и ComSys, Computer Science and Information Systems. Један је од уредника зборника са 29th *International Conference on Lexis and Grammar* и зборника *Workshop on Computational Linguistics and Natural Language Processing of Balkan Languages*. Као позвани професор, држала је предавања у Паризу, на Универзитету Paris-Est, у Институту за језик Бугарске академије наука и на Универзитету Франсоа Рабле у Туру.

Учествовала је као ментор или члан комисије за одбрану великог броја мастер, магистарских и докторских радова на Филолошком факултету, као и на Математичком факултету Универзитета у Београду.

II. БИБЛИОГРАФИЈА

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

(1) Библиографија др Цветане Крстев до избора у звање ванредног професора (до 2008. године)

Књиге и електронска издања објављени до 2008.

1. **Cvetana Krstev** [prevela sa slovenačkog], Ivan Bratko, Vladislav Rajković, *Računarstvo s programskim jezikom paskal*, Nolit, Beograd, 1989.
2. **Cvetana Krstev** [prevela sa engleskog], Kolin A. Dej, *Računarska obrada teksta*, Nolit, Beograd, 1990.
3. **Cvetana Krstev** [računarska obrada leksike], *Vukove narodne poslovice s registrom ključnih reči*, ур. Slobodan Đorđević, Nolit, Beograd, 1996.
4. Duško Vitas, Goran Nenadić, **Cvetana Krstev** [Electronic edition of Serbian translation of Plato's Republic aligned with 17 languages by Duško Vitas, Goran Nenadić, Cvetana Krstev], *East meets West – A compendium of Multilingual Resources*, eds. Tomaž Erjavec, Ann Lawson, Laurent Romary, TELRI Association e.V., Institut für deutsche Sprache, Mannheim, 1998.
5. Duško Vitas, **Cvetana Krstev**; [Electronic edition of Serbian translation of Orwell's 1884 aligned with 7 languages by Duško Vitas, Cvetana Krstev], *East meets West – A compendium of Multilingual Resources*, eds. Tomaž Erjavec, Ann Lawson, Laurent Romary, TELRI Association e.V., Institut für deutsche Sprache, Mannheim, 1998.
6. **Cvetana Krstev**, *Processing of Serbian – Automata, Texts and Electronic dictionaries* Faculty of Philology, University of Belgrade, Belgrade, 2008.

Поглавља у књигама и тематским зборницима до 2008.

1. Duško Vitas, Gordana Pavlović-Lažetić, **Cvetana Krstev**, "Languages of Yugoslavia", in *The Humanities Computing Yearbook 1989-1990*, ed. Ian Lancashire, pp. 313-326, Clarendon Press, Oxford, 1991.
2. Duško Vitas, Gordana Pavlović-Lažetić, **Cvetana Krstev**, "Electronic Dictionary and Text Processing in Serbo-Croatian", in *Sprache – Kommunikation – Informatik: Akten des 26. Linguistischen Kolloquiums*, Poznań 1991, eds. Józef Darski, Zygmunt Vetulani, pp. 225-233, Max Niemezer Verlag, Tübingen, 1993.
3. Duško Vitas, **Cvetana Krstev**, "Some Aspects of Electronic Dictionary Construction", in *Brain and Speech : Neurocomputing and Speech Recognition*, ed. Jelisaveta Milojević, pp. 171-182, Filološki fakultet, Beograd, 1996.
4. **Cvetana Krstev**, Duško Vitas, Gordana Pavlović-Lažetić, "Neutralization of Variations in the Structure of a Dictionary Entry in Serbo-Croatian", in *Formal Slavistik*, eds. Uwe Junghanns, Gerhild Zybatow, pp. 417-426, Vervuert Verlag, Frankfurt am Main, 1997.
5. Duško Vitas, **Cvetana Krstev**, Gordana Pavlović-Lažetić, "Flexible Dictionary Entry", in *Current Issues in Formal Slavic Linguistics*, eds. Gerhild Zybatow, Uwe Junghanns, Grit Mehlhorn, Luka Szucsich, pp. 461-468, Peter Lang, Frankfurt am Main ; Berlin ; Bern ; Bruxelles ; New York ; Oxford ; Wien, 2001.
6. Goran Nenadić, Duško Vitas, **Cvetana Krstev**, "Local Grammars and Compound Verb Lemmatization in Serbo-Croatian", in *Current Issues in Formal Slavic Linguistics*, eds. Gerhild Zybatow, Uwe Junghanns, Grit Mehlhorn, Luka Szucsich, pp. 469-477, Peter Lang, Frankfurt am Main ; Berlin ; Bern ; Bruxelles ; New York ; Oxford ; Wien, 2001.
7. 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.
8. Duško Vitas, Gordana Pavlović-Lažetić, **Cvetana Krstev**, "About Word Length Counting in Serbian", in *Contributions to the Science of Text and Language—Word Length Studies and Related Issues*, editor Peter Grzybek, serija "Text, Speech and Language Technology", volume 31, eds. N. Ide and J. Veronis, pp. 301-318, Springer, Berlin, Heidelberg, 2006.
9. Duško Vitas, **Cvetana Krstev**, "Local Grammars in Word Counting", in *Exact Methods in the Study of Language and Text*, eds. Peter Grzybek and Reinhard Kohler, pp. 665-679, Mouton de Gruyter, Berlin, New York, 2007.
10. **Cvetana Krstev**, Duško Vitas, Gordana Pavlović-Lažetić, "Resources and Methods in the Morphosyntactic Processing of Serbo-Croatian", In *Formal Description of Slavic Languages*, Gerhild Zybatow et al. (eds), Peter Lang: Frankfurt am Main, pp. 3-17, 2008.

Радови у часописима до 2008.

1. **Cvetana Krstev**, "Naučni skup "Računarska obrada jezičkih podataka"", oktobar 1985, Bled, u *INFORMATIKA*, vol. 20, No. 1, pp. 47-49, Beograd, 1986.
2. **Cvetana Krstev**, "Pogled na TeX", u *Računarstvo u nauci i obrazovanju*, vol. 1, No. 2-4, pp. 118-122, Beograd, 1987.
3. Nedeljko Parezanović, Svetozar Niketić, Miomir Todosijević, Duško Vitas, **Cvetana Krstev**, "Računarska laboratorija Prirodno-matematičkog fakulteta - koncepcija, prvi rezultati i perspektive", u *Praksa : jugoslovenska revija za informatiku i AOP*, vol. 23, No. 3-4, pp. 26-28, 1988.

4. **Cvetana Krstev**, Smiljana Jović, Duško Vitas, "Projekat informacionog sistema "Studentska služba Prirodno-matematičkih fakulteta", u *Praksa : jugoslovenska revija za informatiku i AOP*, vol. 3-4, pp. 29-35, 1988.
5. **Cvetana Krstev**, "Serbo-Croatian Hyphenation: a TeX point of view", in *TUGboat*, Providence, Rhode Island, vol. 12, No. 2, pp. 215-223, 1991.
6. **Cvetana Krstev**, "Literary and Linguistic computing in Yugoslavia in 1990", in *Literary and Linguistic Computing*, vol. 6, No. 4, pp. 296-297, Oxford, 1991.
7. **Cvetana Krstev**, "Algoritmi za sravnjivanje niski karaktera", u *Računarstvo*, vol. 2, pp. 25-34, Beograd, 1993.
8. Duško Vitas, **Cvetana Krstev**, "Cultural impacts on electronic publishing: experience in Serbia", in *New Library World*, vol. 100, No. 1149, pp. 171-178, 1999.
9. **Cvetana Krstev**, "Budućnost teksta", in *Svet reči : srednjoškolski časopis za srpski jezik i književnost*, ed. Jovan Deretić, No. 9-12, pp. 85-90, Čigoja štampa, Beograd, 1999-2000.
10. **Cvetana Krstev**, "Formati za predstavljanje informacija: otvorena pitanja", in *INFOteka : časopis za informatiku i bibliotekarstvo*, vol.1, No. 2, pp. 51-57, Beograd, 2000.
11. **Cvetana Krstev**, "Srpski jezik u informatičkom okruženju", in *Književnost i jezik*, No. 1-2, pp. 21-32, Beograd, 2000.
12. **Cvetana Krstev**, "Digitalne biblioteke: razgraničenje pojmova, " in *INFOteka : časopis za informatiku i bibliotekarstvo*, vol. 3, No.1, pp. 3-14, Beograd, 2002.
13. Duško Vitas, **Cvetana Krstev**, Olga Sabo, "O predstavljanju morfološke informacije u elektronskim rečnicima slovenskih jezika", in *Književnost i jezik*, vol. 50, No. 1-3, pp. 97-106, Beograd, 2003.
14. **Cvetana Krstev**, Gordana Pavlović-Lažetić, Ivan Obradovic, Duško Vitas, "Corpora Issues in Validation of Serbian Wordnet", in *Proceedings of the 6th International Conference TSD 2003 : Text, Speech and Dialogue*, Czech Republic, September 8-12, 2003, eds. Vaclav Matoušek, Pavel Mautner, serija "Lecture Notes in Artificial Intelligence" : Subseries of Lecture Notes in Computer Science, eds. J.G. Carbonell, J. Siekmann, pp. 132-137, Springer, Berlin, Heidelberg, 2003.
15. Gordana Pavlović-Lažetić, Duško Vitas, **Cvetana Krstev**, "Towards Full Lexical Recognition", in *Proceedings of the 7th International Conference TSD 2004 : Text, Speech and Dialogue*, Brno, Czech Republic, September 8-11, 2004, eds. Petr Sojka, Ivan Kopček, Karel Pala, serija "Lecture Notes in Artificial Intelligence" : Subseries of Lecture Notes in Computer Science, eds. J.G. Carbonell, J. Siekmann, pp. 179-186, Springer, Berlin, Heidelberg, 2004.
16. **Cvetana Krstev**, Gordana Pavlović-Lažetić, Duško Vitas, Ivan Obradović, "Using Textual and Lexical Resources in Developing Serbian Wordnet", in *Romanian Journal of Information Science and Technology*, vol. 7, No. 1-2, pp. 147-161, Romanian Academy, Publishing House of the Romanian Academy, 2004.
17. **Cvetana Krstev**, Duško Vitas, Tomaz Erjavec, "Morpho-Syntactic Descriptions in MULTTEXT-East - the Case of Serbian", in *Informatica*, No. 28, pp. 431-436, The Slovene Society Informatika, Ljubljana, 2004.
18. Duško Vitas, **Cvetana Krstev** (2005), "Regular derivation and synonymy in an e-dictionary of Serbian", in *Archives of Control Sciences*, Volume 51(LI), No. 3, pp. 469-480, Committee of Automation and Robotics, Polish Academy of Sciences, 2005.
19. **Cvetana Krstev**, Duško Vitas, Agata Savary, "Prerequisites for a Comprehensive Dictionary of Serbian", in *Proceedings of the 5th International Conference on NLP, FinTAL 2006*, Turku, Finland, August, 2006, eds. Tapio Salakoski, Filip Ginter, Sampo Pyysalo, Tapio Pahikkala, serija "Lecture Notes in Artificial Intelligence" : Subseries of

- Lecture Notes in Computer Science, eds. J.G. Carbonell, J. Siekmann, pp. 552-564, Springer, Berlin, Heidelberg, 2006.
20. Duško Vitas, **Cvetana Krstev**, Eric Laporte, "Preparation and exploitation of Bilingual Texts", in *Lux Coreana*, No. 1, pp. 110-132, Han-Seine, 2006.
 21. Duško Vitas, **Cvetana Krstev**, Denis Maurel "A note on the Semantic and Morphological Properties of Proper Names in the Prolex Project", in *Lingvisticae Investigationes*, Special issue on Named Entities: Recognition, Classification and Use, eds. Satoshi Sekine and Elisabete Ranchhod, Vol. 30(XXX), No. 1, pp. 115-134, John Benjamins Publishing Company, Amsterdam, Philadelphia, 2007.
 22. Denis Maurel, Duško Vitas, **Cvetana Krstev**, Svetla Koeva, "Prolex: a lexical model for translation of proper names. Application to French, Serbian and Bulgarian", in *Bulag - Bulletin de Linguistique Appliquée et Générale*, Les langues slaves et le français : approches formelles dans les études contrastives, eds. Aleksandra Dziadkiewicz et Izabella Thomas, No. 32, pp. 55-72, Presses Universitaires de Franche Comté, Besancon, 2007.
 23. Agata Savary, **Cvetana Krstev**, Duško Vitas, "Inflectional non compositionality and variation of compounds in French, Polish and Serbian, and their automatic processing", in *Bulag - Bulletin de Linguistique Appliquée et Générale*, Les langues slaves et le français : approches formelles dans les études contrastives, eds. Aleksandra Dziadkiewicz et Izabella Thomas, No. 32, pp. 73-94, Presses Universitaires de Franche Comté, Besancon, 2007.
 24. **Cvetana Krstev**, Sandra Gucul "Ka dečjem rečniku - automatska obrada dečje literature", u *Bibliotekar : časopis za teoriju i praksu bibliotekarstva*, god. XLIX, sv. 3-4, pp. 295-310, Beograd, 2007.

Радови у зборницима са конференција до 2008.

1. **Cvetana Krstev**, "Rastavljanje reči srpskohrvatskog jezika na kraju retka", u *Zbornik radova sa III naučnog skupa "Računarska obrada jezičkih podataka"*, Bled, oktobar 1985., eds. Damjan Bojadžijev, Petar Tancig, Duško Vitas, pp. 289-301, Institut "Jožef Štefan", Društvo za uporabno jezikoslovje Slovenije, Ljubljana, 1985.
2. **Cvetana Krstev**, "Prikaz sistema TeX i METAFONT D.E. Knutha", u *Zbornik radova simpozijuma INTERGRAFIKA*, ed. Nikola Barišić, pp. 51-64, Viša grafička škola, Zagreb, 1985.
3. **Cvetana Krstev**, "O problemu primene i adaptiranja programske opreme za obradu teksta u neengleskim govornim sredinama", u *Zbornik radova sa IV naučnog skupa "Računarska obrada jezičkih podataka"*, Portorož, 3-7 oktobar 1988., eds. Damjan Bojadžijev, Petar Tancig, Duško Vitas, pp. 289-301, Institut "Jožef Štefan", Društvo za uporabno jezikoslovje Slovenije, Ljubljana, 1988.
4. **Cvetana Krstev**, Smiljana Jović-Puač i Duško Vitas, "Analiza podjezika uputstava za lekove na srpskohrvatskom i slovenačkom jeziku", deo I, u *Zbornik radova sa IV naučnog skupa "Računarska obrada jezičkih podataka"*, Portorož, 3-7 oktobar 1988., eds. Damjan Bojadžijev, Petar Tancig, Duško Vitas, pp. 249-255, Institut "Jožef Štefan", Društvo za uporabno jezikoslovje Slovenije, Ljubljana, 1988.
5. Vesna Vučković, **Cvetana Krstev**, Ilija Lalović, "Uporedna analiza programskih paketa za izradu konkordanci", u *Zbornik radova sa IV naučnog skupa "Računarska obrada jezičkih podataka"*, Portorož, 3-7 oktobar 1988., eds. Damjan Bojadžijev, Petar Tancig, Duško Vitas, pp. 245-248, Institut "Jožef Štefan", Društvo za uporabno jezikoslovje Slovenije, Ljubljana, 1988.
6. Duško Vitas, **Cvetana Krstev**, "On representing morpho-graphemic information in a Serbo-Croatian Dictionary", in *Međunarodni seminar po mašnomu prevodu "IVM i*

- perevod 89*”, Tbilisi, 27 novembar - 2 decembar, 1989 g., eds. I.I. Ubin, E.E. Lovckiy, pp. 332-335, Vsesoznij centr perevodov naučno-tehničkoj literaturi i dokumentacij ; Gruzinskij naučno-issledovateljskij institut naučno-tehničkoj informacii, Moskva, 1988.
7. **Cvetana Krstev**, “O izboru programskog jezika za obradu jezičkih podataka”, u *Zbornik radova sa simpozijuma “Matematička i računarska lingvistika - teorija i praksa”*, Beograd, 13-14 XI 1987., ed. Duško Vitas, pp. 128-133, Društvo za primenjenu lingvistiku Srbije, Beograd, 1990.
 8. Duško Vitas, **Cvetana Krstev**, Smiljana Jović-Puač, “Računarska lingvistika u Računarskoj laboratoriji Prirodno-matematičkih fakulteta u Beogradu”, u *Zbornik radova sa simpozijuma “Matematička i računarska lingvistika - teorija i praksa”*, Beograd, 13-14 XI 1987., ed. Duško Vitas, pp. 149-152, Društvo za primenjenu lingvistiku Srbije, Beograd, 1990.
 9. **Cvetana Krstev**, “Standardi za elektronsko rukovanje dokumentima”, u *Savetovanje “Statistika i informaciona tehnologija”*, Arandelovac, 29-30. maja 1990., pp. 133-140, Statističko društvo Srbije, Republički zavod za statistiku SR Srbije, Beograd, 1990.
 10. **Cvetana Krstev**, “Standardi za pripremu tekstova u mašinski čitljivom obliku,” in *7. jugoslovenska konferencija “Standardizacija za Evropu 1992”*, Novi Sad, 5-6. novembar 1990., vol. 1-6, pp. 19, Jugoslovenska organizacija za standardizaciju i kvalitet, Savezni zavod za standardizaciju, Novi Sad, 1990.
 11. Duško Vitas, **Cvetana Krstev**, “Interaction between Dictionary and Text in Serbo-Croatian”, in *Papers in Computational Lexicography COMPLEX 1992*, Budapest, 4-8 October 1992, eds. Ferenc Kiefer, Gábor Kiss, Júlia Pajzs, pp. 333-342, Linguistic Institute, Hungarian Academy of Sciences, Budapest, 1992.
 12. **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.
 13. **Cvetana Krstev**, “Standardi za opis strukture i izgleda teksta i hiperteksta”, u *Zbornik radova sa savetovanja “Standardizacija i kvalitet u informacionim tehnologijama”*, Beograd, 1. i 2. juni 1995., pp. 1-5, Savezno ministarstvo za razvoj, nauku i životnu sredinu, Agencija BBN Congress Management, Beograd, 1995.
 14. **Cvetana Krstev**, Duško Vitas, Henning Moerk, Gordana Pavlović-Lažetić, “Graphic or Graphemic Encoding of Medieval Slavic Manuscripts”, in *Book of abstracts of the First International Conference “Computer Processing of Medieval Slavic Manuscripts”*, Blagoevgrad, 24-29 July 1995, eds. David J. Birnbaum, Andrej T. Bojadžiev, Milena P. Dobрева, Anisava L. Miltenova, pp. 17-19, Institute of Literature, Bulgarian Academy of Sciences, Sofia, 1995.
 15. Duško Vitas, **Cvetana Krstev**, Gordana Pavlović-Lažetić, “Dictionary Layers Underlying Electronic Texts”, in *ZUMA-Nachrichten Spezial “Text Analysis and Computers”*, Mannheim, 18-21 September 1995, eds. Cornelia Zuell, Janet Harkness, Juergen H.P. Hoffmeyer-Zlotnik, pp. 131, Zentrum für Umfragen, Methoden und Analysen (ZUMA), Mannheim, 1996.
 16. Duško Vitas, **Cvetana Krstev**, “Tuning the Text with an Electronic Dictionary”, in *Papers in Computational Lexicography COMPLEX 1996*, Budapest, 15-17 September 1996, eds. Ferenc Kiefer, Gábor Kiss, Júlia Pajzs, pp. 267-276, Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest, 1996.
 17. **Cvetana Krstev**, “An Approach to the Text Structure Description”, u *Zbornik sa VI naučnog skupa o sistemu naučnih i tehnoloških informacija “Mreže specijalizovanih*

sistema naučnih, tehnoloških i poslovnih informacija u SR Jugoslaviji – stanje i perspektive”, Kotor, 19-21 jun, 1996.

18. Duško Vitas, **Cvetana Krstev**, Gordana Pavlović-Lažetić, “Analiza jugoslovenskih terminoloških standarda sa područja informatike”, u *Zbornik radova sa jugoslovenske konferencije “Standardizacija terminologije”*, Beograd, 10-12 juni, 1996., pp. 149-154, 1996.
19. **Cvetana Krstev**, Duško Vitas, “Morphological Normalization of Translation Equivalents”, in *Proceedings of the Third European Seminar “Translation Equivalence”*, Montecatini Terme, Italy, October 16-18, 1997, eds. Wolfgang Teubert, Elena Tognini Bonelli, Norbert Volz, pp. 117-124, The TELRI Association e.V., Institut für deutsche Sprache, Montecatini Terme, Tuscan Word Center, Mannheim, 1998.
20. Duško Vitas, Cvetana Krstev, Gordana Pavlović-Lažetić, Goran Nenadić, “Recent Results in Serbian Computational Lexicography”, in *Proceedings of the Symposium “Contemporary Mathematics”*, Belgrade, 18-20 December 1998, ed. Neda Bokan, pp. 113-130, Faculty of Mathematics, Belgrade, 2000.
21. Cvetana Krstev, “Information Science Curriculum at the Undergraduate Studies of Library and Information Science”, in *Skup bibliotekara balkanskih zemalja : Saradnja obrazovanje kvalitet*, eds. Aleksandra Vraneš, Ljiljana Marković, pp. 117-122, Narodna biblioteka Srbije, Beograd, 2002.
22. Duško Vitas, Cvetana Krstev, “Composite Tense Recognition and Tagging in Serbian”, in *Proceedings of the Workshop on Morphological Processing of Slavic Languages : 10th Conference of the European Chapter, EACL 2003*, Budapest, Hungary, April 13th, 2003, eds. Tomaž Erjavec and Duško Vitas, pp. 54-61.
23. Tomaž Erjavec, Cvetana Krstev, Vladimir Petkevič, Kiril Simov, Marko Tadić, Duško Vitas, “The MULTEXT-East Morphosyntactic Specifications for Slavic Languages”, in *Proceedings of the Workshop on Morphological Processing of Slavic Languages : 10th Conference of the European Chapter, EACL 2003*, Budapest, Hungary, April 13th, 2003, eds. Tomaž Erjavec and Duško Vitas, pp. 25-32.
24. Ivan Obradović, Cvetana Krstev, Gordana Pavlović-Lažetić, Duško Vitas, “Corpus Based Validation of WordNet Using Frequency Parameters”, in *Proceedings of the GWC : Second International WordNet Conference*, Brno, Czech Republic, January 20-23, 2004, eds. P. Sojka, K. Pala, P. Smrž, Ch. Fellbaum, P. Vossen, ed. 1, pp. 181-186, Masaryk University, Brno, 2004.
25. Cvetana Krstev, Duško Vitas, Tomaz Erjavec, “MULTEXT-East Resources for Serbian”, in *Zbornik 7. mednarodne multikonference “Informacijska družba IS 2004” Jezikovne tehnologije 9-15 Oktober 2004*, Ljubljana, Slovenija, eds. Tomaz Erjavec, Jerneja Zganec Gros, Institut “Jozef Stefan”, Ljubljana, 2004.
26. Ranka Stanković, Cvetana Krstev, Duško Vitas, Ivan Obradović, Gordana Pavlović-Lažetić “Integriranje heterogenih lekisičkih resursa”, in *Festivalski katalog 11. Festivala informatičkih dostignuća INFOFEST 2004*, 26th September - 2nd October, 2004, Budva, Serbia & Montenegro, pp. 308-316, 2004.
27. Duško Vitas, Cvetana Krstev, “Derivational Morphology in an E-Dictionary of Serbian”, in *Proceedings of 2nd Language & Technology Conference*, April 21-23, 2005, Poznań, Poland, ed. Zygmunt Vetulani, pp. 139-143, Wydawnictwo Poznańskie Sp. z o.o., Poznań, 2005.
28. Svetla Koeva, Cvetana Krstev, Ivan Obradović, Duško Vitas, “Resources for Processing Bulgarian and Serbian - a brief overview of Completeness, Compatibility and Similarities”, in *Workshop on Language and Speech Infrastructure for Information Access in the Balkanic Countries*, 25 September 2005, Borovets, Bulgaria, eds. S. Piperidis and E. Paskaleva, pp. 31-38, 2005.

29. Mickaël Tran, Denis Maurel, Duško Vitas, Cvetana Krstev, "A French-Serbian Web Collaborative Work on a Multilingual Dictionary of Proper Names", Papillon 2005 workshop on Multilingual Lexical Databases, in Association with the Sixth Symposium on Natural Language Processing (SNLP 2005), Chiang Rai, Thailande, 12-14 décembre, 2005.
30. **Cvetana Krstev**, "U potrazi za skrivenom informacijom", u *Ekonomska uloga biblioteka u savremenom društvu : zbornik radova sa međunarodnog naučnog skupa održanog u Beogradu od 7. do 9. oktobra 2004. godine / urednici Aleksandra Vraneš, Ljiljana Marković*. Beograd : Filološki fakultet Univerziteta u Beogradu : Fakultet za menadžment Univerziteta "Braća Karić" : Bibliotekarsko društvo Srbije, 2005 (Beograd : MST Gajić), str. 167 (sažetak)
31. **Cvetana Krstev**, Ranka Stanković, Duško Vitas, Ivan Obradović, "WS4LR - a Workstation for Lexical Resources", in *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, Genoa, Italy, May 2006, pp. 1692-1697, 2006.
32. **Cvetana Krstev**, Ivan Obradović, Duško Vitas, "Developing Balkan specific concepts within BalkaNet - a multilingual database of semantic networks", in *Proceedings of the 5th International Conference Formal Approaches to South Slavic and Balkan Languages*, FASSBL 2006, Sofia, Bulgaria, October, 2006, eds. Svetla Koeva and Mila Dimitrova-Vulchanova, pp. 94-98, The Department of Computational Linguistics, Institute of Bulgarian Language, 2006.
33. **Cvetana Krstev**, Duško Vitas, Vesna Šatev, "TEI Encoding of Serbian Proverbs", in *Computer Applications in Slavic Studies*, *Proceedings of Azbuky.Net International Conference and Workshop*, 24-27 October 2005, Sofia, Bulgaria, eds. Anisava Miltenova et al, pp. 101-114, Institute of Literature, Bulgarian Academy of Sciences, Sofia, Bulgaria, 2006.
34. **Cvetana Krstev**, "Specifični koncepti Balkana u semantičkoj mreži Wordnet", u *Zborniku radova "Susreti kultura"*, Novi Sad, decembar 2004, eds. Ljiljana Subotić et al, pp. 275-285, Univerzitet u Novom Sadu, Filozofski fakultet, Novi Sad, 2006.
35. Duško Vitas, **Cvetana Krstev**, Svetla Koeva, "Towards a Complex Model for Morpho-Syntactic Annotation", in *Proceedings of the Workshop Workshop on a Common Natural Language Processing Paradigm for Balkan Languages*, 26 September 2007, Borovets, Bulgaria, eds. E. Paskaleva and Milena Slavcheva, pp. 65-71, 2007.
36. **Cvetana Krstev**, Duško Vitas, "Treatment of Numerals in Text Processing", in *Proceedings of 3rd Language & Technology Conference*, October 5-7, 2007, Poznań, Poland, ed. Zygmunt Vetulani, pp. 418-422, IMPRESJA Wydawnictwa Elektroniczne S.A., Poznań, 2007.
37. Svetla Koeva, **Cvetana Krstev**, Duško Vitas, "Morpho-semantic Relations in WordNet - a Case Study for two Slavic Languages", In the *Proceedings of Global WordNet Conference 2008*, eds. Attila Tanacs et al, University of Szeged, Department of Informatics, pp. 239-253, 2007.
38. **Cvetana Krstev**, „Digitalna era i biblioteke : tekstovi u digitalnom obliku - više od štampanja“, u *Savremena školska biblioteka : zbornik predavanja održanih 10. i 11. novembra na Filološkom fakultetu u Beogradu / urednik Aleksandra Vraneš*. Beograd : Bibliotekarsko društvo Srbije, 2007 (Beograd : Belpak), str. 51 (sažetak).
39. **Cvetana Krstev**, „Internet kao izvor informacija“, u *Školski bibliotekar saradnik u nastavi / urednici Aleksandra Vraneš, Ljiljana Marković, Predrag Stanojević*. Beograd : Filološki fakultet Univerziteta : Bibliotekarsko društvo Srbije, 2008 (Beograd : Belpak), str. 73-88.

Учешће на домаћим научним пројектима до 2008. године

1. 1981-1985: Научни пројекат 1.35 „Математичка и рачунарска лингвистика“, финансирао РЗН СР Србије (руководиоци проф. Душан Јовић, проф. Марица Прешић);
2. 1983-1985: Развојни пројекат „Аутоматска обрада текста“, финансирао ОЗН Београда;
3. 1985-1986: Развојни пројекат „Адаптација ТИТ-тезауруса са енглеског на српскохрватски језик“, за потребе Југословенског завода за стандардизацију, финансирао Савезни секретаријат за науку, Београд;
4. 1986-1990: Научно-истраживачки пројекат 1.51 „Рачунарство са применама“ - финансира РЗН СРС;
5. 1985-1988: Развојни пројекат Т-135.6.2. „Јединствени информациони систем Београдског универзитета - подсистем: Студирање“, координатор институт Михаило Пупин, финансирао ОЗН Београда и РЗН СРС;
6. 1986-1990: Научно-истраживачки пројекат Т-142 „Вештачка интелигенција са применама“, тема „Рачунарско разумевање природног језика“, финансира РЗН СРС, координатор институт Михаило Пупин;
7. 1988: Развојни пројекат „Дугорочни програм стандардизације на подручју информатике“, финансирао Савезни комитет за науку и Југословенски завод за стандардизацију, Београд;
8. 1989-1996: Научно-истраживачки пројекат П-84 „Вештачка интелигенција“ потпројекат „Обрада српскохрватског“, финансирао Савезни комитет за развој;
9. 1996-2000: учесник на пројекту 04М02 „Математичка логика, комбинаторика и обрада информација“;
10. 2002-2005: Научно-истраживачки пројекат 1743 „Интеракција текста и речника“, учесници: Математички факултет, Филолошки факултет, Рударско-геолошки факултет (Београд) и Филозофски факултет (Нови Сад);
11. 2006-2010: Научно-истраживачки пројекат 148921А „Теоријско-методолошки оквир за модернизацију описа српског језика“, учесници: Математички факултет, Филолошки факултет, Рударско-геолошки факултет (Београд).

Учешће на међународним научним пројектима до 2008. године

1. 1983-1986: Кооперативни научни пројекат „Аутоматска обрада текстуелних информација на језицима југословенских народа“, координатор СРЦЕ, Загреб, учесници: СРЦЕ, Математички институт, Београд и институт "Јошеф Стефан", Љубљана, реализован у оквиру размене СФРЈ са САД;
2. 1989-1991: Кооперативни пројекат „Languages Industries“, финансирао Европски савет;
3. 1995-2001: Европски истраживачки пројекат TELRI 1 и TELRI 2 (Trans-European Language Resources Infrastructure), финансиран у оквиру програма Copernicus, (Copernicus Concerted Action 1202), координатор Институт за немачки језик у Манхајму;
4. 2002-2004: Европски истраживачки пројекат „BalkaNet - Design and Development of a Multilingual Balkan WordNet“, финансира Европска комисија у оквиру програма FW5, координатор Универзитет у Патрасу, Грчка;
5. 2004-2005: Научно-истраживачки пројекат „Вишејезични речник властитих имена“, пројекат мобилности у оквиру програма „Павле Савић“ научне сарадње са Француском, француски партнер: prof. Denis Maurel, Université Tours;

6. 2004-2005: Научно-истраживачки пројекат „Ресурси за аутоматско преводјење између српског и словеначког“, пројекат мобилности у оквиру програма научне сарадње са Словенијом, словеначки партнер: др Томаж Ерјавец, Институт „Јожеф Стефан“, Љубљана;
7. 2007-2008: SEE-ERA.NET - Building Language Resources and Translation Models for Machine Translation focused on South Slavic and Balkan Languages (ICT 10503 RP);
8. 2007 - 2008: WISE - *An Electronic Marketplace to Support Pairs of Less Widely Studied European Languages* (Електронско тржиште за подршку паровима мање изучаваних европских језика) (BSEC 009 / 05.2007), координатор Универзитет у Патрасу.

Учешће у комисијама за одбрану магистарских и докторских радова до 2008. године

Докторске тезе

1. Здравко Вукчевић: *Информативна структура Његошевог пјесничког језика*, коментор, Филолошки факултет, 2003.

Магистарске тезе

1. Бранкица Поткоњак-Лукић: *Контролисани речник појмова у области војних наука*, ментор, Филолошки факултет, 2005;
2. Даница Радовановић: *Електронско издаваштво у академском окружењу – нове форме и изазови*, ментор, Филолошки факултет, 2006;
3. Драгана Димитријевић-Сабовљев: *Аутоматизација процеса пописивања грађе и формирања електронског записа у припреми израде библиографије*, ментор, Филолошки факултет, 2007.

Језички ресурси развијени до 2008. године

1. *Програм за растављање српскохрватских речи на слоге* - Систем образаца развијен за TeX и LaTeX, 1989;
2. *Пројектовање информационог система Београдског универзитета - подсистем студирање*, Београдски универзитет, РЗН СРС и ОЗН Београда, систем развијан од 1986 – 1989;
3. *Први рачунарски корпус савременог српскохрватског језика* (око милион речи везаног текста на манетном медијуму) настао у периоду од 1981. - 1990, намењен статистичкој анализи података о језику;
4. Развој система *LeXimir* (раније *WS4LR*), радне станице за руковање језичким ресурсима (од 2006. – систем се развија и данас);
5. *Паралелни српско-енглески корпус* за потребе пројекта INTERA (ELDA: <http://www.elda.org/intera>). Морфосинтаксички анотиран и лематизовани паралелизовани корпус у укупној дужини од 2 милиона текућих речи (оба језика);
6. Вишејезична онтологија властитих имена PROLEX (у сарадњи са Универзитетом у Туру);
7. *Морфосинтаксички опис српског у формату MULTEXT-East v.3* (доступан уз лиценцу на <http://nl.ijs.si/ME/V3/>);
8. Морфолошки електронски речник српског језика у LADL-формату (са Д. Витасом) са око 130.000 одредница (обим једнотомног речника савременог српског језика),

са флективним, деривационим и синтаксичко-семантичким информацијама – развија се даље и данас;

9. Семантичка мрежа типа WordNet (са коауторима) са око 21.000 синсетова на српском и припадајућим семантичким релацијама, повезана са другим европским језицима – развија се даље и данас.

(2) Библиографија др Цветане Крстев од избора у звање ванредног професора (од 2008. године)

(2.а) Монографије и поглавља у монографијама

- 2.1 Duško Vitas, **Cvetana Krstev**, “O paralelnim korpusima, a posebno o beogradskim paralelnim korpusima i načinu njihove eksploatacije”, Sonderdruck aus: Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen (Razlike između bosanskog/bošnjačkog, hrvatskog i srpskog jezika), ed. Branko Tošović, pp. 630-649, Münster, Germany, LITVerlag, 2008.
- 2.2 **Cvetana Krstev**, “Digitalne biblioteke kao potencijalni lingvistički resurs - stanje u Srbiji”, In: *Serbian Linguistics: Setting the Agenda*, eds B. Golubović & C. Voß . SLCCEE, Bd. 7, pp. 123-142, Verlag Otto Sagner: München, 2009.
- 2.3 Duško Vitas, **Cvetana Krstev**, Denis Maurel, “A note on the semantic and morphological properties of proper names in the Prolex project”, In: *Named Entities - Recognition, classification and use*, eds. S. Sekine and E. Rasmann, Benjamins Current Topics 19, pp. 117-136, John Benjamins Publishing Company, ISBN 978-90-272-2249-7, 2009.
- 2.4 **Cvetana Krstev**, Ranka Stanković, Ivan Obradović, Duško Vitas, Miloš Utvić, “Automatic Construction of a Morphological Dictionary of Multi-Word Units”, in *Proceedings of the 7th International Conference on NLP, IceTAL 2010*, Reykjavik, Iceland, August 16-18, 2010. eds. Hrafn Loftsson, Eiríkur Rögnvaldsson, Sigrún Helgadóttir, Lecture Notes in Computer Science 6233 Springer 2010, ISBN 978-3-642-14769-2, pp. 226-237, 2010.
- 2.5 Ivan Obradović, **Cvetana Krstev**, Duško Vitas, “Measuring semantic relevance of words in synsets”, In: *Text and Language. Structures · Functions · Interrelations. Quantitative Perspectives*, eds. Peter Grzybek, Emmerich Kelih and Ján Mačutek, pp. 133-144, Praesens Verlag, Wien, ISBN 978-3-7069-0625-8, 2010.
- 2.6 Ivan Obradović, Aljoša Obuljen, Duško Vitas, **Cvetana Krstev**, Vanja Radulović, “Distribution of canonical syllable types in Serbian”, In: *Text and Language. Structures · Functions · Interrelations. Quantitative Perspectives*, eds. Peter Grzybek, Emmerich Kelih and Ján Mačutek, pp. 145-157, Praesens Verlag, Wien, ISBN 978-3-7069-0625-8, 2010.
- 2.7 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.
- 2.8 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.
- 2.9 Duško Vitas, **Cvetana Krstev**, “Les méthodes d'alignement de textes pour le développement d'un corpus numériques judéo-espagnol”, in *Recensement, analyse et*

traitement numérique des sources écrites pour les études séfardes, eds. Soufiane Rouissi et Ana Štulić-Etchevers, pp. 327-342, Presses Universitaires de Bordeaux, Collection de la Maison des pays ibériques, série littéralité, Bordeaux, 2013, ISBN 978-2-86781-813-4, ISSN 0296-7588

- 2.10 **Cvetana Krstev**, Ivan Obradović, Ranka Stanković, Duško Vitas, “An Approach to Efficient Processing of Multi-word Units”, in *Computational Linguistics - Applications*, eds. Adam Przepiórkowski, Maciej Piasecki, Krzysztof Jassem, Piotr Fuglewicz, Studies in Computational Intelligence 458 (Series ISSN 1860-949X), Springer-Verlag, Berlin Heidelberg, ISBN 978-3-642-34398-8 (online 978-3-642-34399-5), DOI 10.1007/978-3-642-34399-5_6, pp. 109-229, 2013.

(2.6) Уредништво

- 2.11 *Proceedings of the 29th International Conference on Lexis and Grammar - LGC 2010*, September 15-18th, 2010, Belgrade / edited by Duško Vitas and Cvetana Krstev, Belgrade : Faculty of Mathematics, 2010; ISBN 978-86-7589-080-5, COBISS.SR-ID 177981708
- 2.12 *Proceedings of the CLoBL 2012: Workshop on Computational Linguistics and Natural Language Processing of Balkan Languages*, 2012 / edited by Cvetana Krstev and Katerina Zdravkova

(2.в) Радови у часописима и зборницима са научних скупова од 2008. године

- 2.13 **Cvetana Krstev**, Ivan Obradović, Duško Vitas, “An Approach to the Development of Language Specific Concepts in Wordnets”, In *Southern Journal of Linguistics*, Special Theme: South Slavic and Balkan Languages, Mila Dimitrova-Vulchanova (ed.), Vo. 29, No. 1/2, pp. 106-118, Department of Modern Linguistics, University of Mississippi, 2008.
- 2.14 **Cvetana Krstev**, Bojana Đorđević, Sanja Antonić, Nevena Ivković-Berček, Zorica Zorica, Vesna Crnogorac, Ljiljana Macura, “Kooperativan rad na dogradnji srpskog wordneta,” in *INFOteka : časopis za informatiku i bibliotekarstvo*, vol. IX, No.1-2, pp. 57-75, Beograd, 2008.
- 2.15 **Cvetana Krstev**, Ranka Stanković, Duško Vitas, Ivan Obradović, “The Usage of Various Lexical Resources and Tools to Improve the Performance of Web Search Engines”, in *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, 28-30 May 2008, European Language Resources Association (ELRA), 2008.
- 2.16 **Cvetana Krstev**, Svetla Koeva, Duško Vitas, “A Dictionary-based Model for Morpho-Syntactic Annotation”, in *Proceedings of the 2nd Linguistic Annotation Workshop*, in scope of the Sixth International Conference on Language Resources and Evaluation (LREC'08), Marrakech, Morocco, 26 May 2008, European Language Resources Association (ELRA), 2008.
- 2.17 Duško Vitas, Svetla Koeva, **Cvetana Krstev**, Ivan Obradović, “*Tour du monde through the dictionaries*”, *Actes du 27eme Colloque International sur le Lexique et la Grammaire*, L'Aquila, 10-13 septembre 2008, eds. M. Constant, T. Nakamura, M. De Gioia, S. Vecchiato, pp.249-256, Université Paris-Est, Institut Gaspard-Monge, 2008.
- 2.18 **Cvetana Krstev**, Duško Vitas, “Informatički pogled na bibliografiju”, u *Srpska bibliografija danas*, ur. Aleksandra Vraneš, pp. 229-241, Matica srpska, Novi Sad, 2008.
- 2.19 Dan Tufiş, Svetla Koeva, Tomaž Erjavec, Maria Gavrilidou, and **Cvetana Krstev**. "Building Language Resources and Translation Models for Machine Translation focused

- on South Slavic and Balkan Languages". In Marko Tadić, Mila Dimitrova-Vulchanova and Svetla Koeva (eds.) *Proceedings of the Sixth International Conference Formal Approaches to South Slavic and Balkan Languages (FASSBL 2008)*, pp. 145-152, Dubrovnik, Croatia, September 25-28, 2008. ISBN 978-953-55375-0-2.
- 2.20 Sandra Gucul-Milojević, Vanja Radulović, and **Cvetana Krstev**. "Usage of NooJ Graphs and Annotation for Information Extraction". In Xavier Blanco and Max Silberztein (eds.) *Proceedings of the 2007 International NooJ Conference*, pp. 103-120, Cambridge Scholars Publishing, 2008. ISBN (13) 978-1-4438-0053-2.
 - 2.21 Ranka Stanković, Duško Vitas, and **Cvetana Krstev**. "The NooJ System as Module within an Integrated Language Processing Environment". In Xavier Blanco and Max Silberztein (eds.) *Proceedings of the 2007 International NooJ Conference*, pp. 228-248, Cambridge Scholars Publishing, 2008. ISBN (13) 978-1-4438-0053-2.
 - 2.22 Sandra Gucul-Milojević, Vanja Radulović, Branka Rašić i **Cvetana Krstev**. "Ponovo o formatu digitalnih tekstova". U Aleksandra Vraneš i Ljiljana Marković (ur.) *Elektronska biblioteka - Međunarodna naučna konferencija, Beograd 25-28. IX 2008*, pp. 323-334, Filološki fakultet Univerziteta u Beogradu, 2008. ISBN (13) 978-86-86419-70-5.
 - 2.23 Sanja Antonić i **Cvetana Krstev**. "Serbian Wordnet for Biomedical Sciences". In *INFORUM 2008: 14th Conference on Professional Information Resources*, Prague, May 28-30, 2008.
 - 2.24 Duško Vitas, **Cvetana Krstev**. "Srpski jezik i SNTPI". U Đuro Kutlača (ur.) *SNTPI '09 - Naučno-stručni skup Sistem naučnih, tehnoloških i poslovnih informacija, Beograd 19. i 20. jun 2009*, pp. 87-90, Fakultet informacionih tehnologija, 2009. ISBN 987-86-912685-0-3.
 - 2.25 Ranka Stanković, Ivan Obradović, **Cvetana Krstev**. "Proširivanje upita zasnovano na leksičkim resursima". U Đuro Kutlača (ur.) *SNTPI '09 - Naučno-stručni skup Sistem naučnih, tehnoloških i poslovnih informacija, Beograd 19. i 20. jun 2009*, pp. 77-80, Fakultet informacionih tehnologija, 2009. ISBN 987-86-912685-0-3.
 - 2.26 **Cvetana Krstev**, Ranka Stanković, Duško Vitas, Svetla Koeva, "E-Connecting Balkan Languages", in *Proceedings of the Workshop Workshop on Multilingual resources, technologies and evaluation for Central and Eastern European Languages*, 17 September 2009, Borovets, Bulgaria, eds. C. Vertan, S. Piperidis, E. Paskaleva and Milena Slavcheva, pp. 23-29, 2009. ISBN 978-954-452-008-3.
 - 2.27 **Cvetana Krstev**, Duško Vitas, "An Effective Methode for Developing a Comprehensive Morphological E-Dicitionary of Compounds", The 28th Conference on Lexis and Grammar, Bergen, 29th September - 3rd October 2009, In *Arena Romanistica* eds. B. Lamiroy, E. Laporte, T. Kyriakopoulou, pp. 204-212, University of Bergen, Department of Foreign Languages, 2009.
 - 2.28 Emeline Lecuit, Denis Maurel, Duško Vitas, **Cvetana Krstev**, "Temporal Expressions: Comparisons in a Multilingual Corpus", in *Proceedings of 4th Language & Technology Conference*, November 6-8, 2009, Poznań, Poland, ed. Zygmunt Vetulani, IMPRESJA Wydawnictwa Elektroniczne S.A., Poznań, 2009.
 - 2.29 Dan Tufiş, Svetla Koeva, Tomaž Erjavec, Maria Gavrilidou, **Cvetana Krstev**, "Building Language Resources and Translation Models for Machine Translation Focused on South Slavic and Balkan Languages", In Jana Macháčová, Katarina Rohsmann (eds.) *Scientific results of the SEE-ERA.NET Pilot Joint Call*, pp. 37-48, Centre for Social Innovation Vienna, Austria, June 2009, ISBN 978-3-200-01567-8.
 - 2.30 **Cvetana Krstev**, Ranka Stanković, Duško Vitas, "A Description of Morphological Features of Serbian: a Revision using Feature System Declaration", *Proceedings of the 7th International Conference on Language Resources and Evaluation*, LREC 2010,

- Valetta, Malta, May 2010, pp.816-819, 2010, eds. Nicoletta Calyolari *et al.*, ISBN 2-9517408-6-7.
- 2.31 Miloš Utvić, Ivan Obradović, **Cvetana Krstev**, Duško Vitas, “The Effects of Multi-Word Tagging on Text Disambiguation”, In: *Proceedings of the 29th International Conference on Lexis and Grammar, September 15-18, 2010, Belgrade, Serbia*, eds. Duško Vitas and Cvetana Krstev, pp. 333-342, Faculty of Mathematics, University of Belgrade, Serbia, ISBN 978-86-7589-080-5, 2010.
 - 2.32 **Cvetana Krstev**, “Uloga i mesto slobodnog softvera u bibliotekama i javnom sektoru”, In: *Zbornik radova sa međunarodne naučne konferencije “Etika u nauci i kulturi”, September 25-27, 2009, Beograd, Srbija*, ur. Aleksandra Vraneš i Ljiljana Marković, pp. 233-247, Filološki fakultet, Univerzitet u Beogradu, ISBN 978-86-86419-97-2, 2010.
 - 2.33 Sandra Gucul-Milojević, Vanja Radulović, and **Cvetana Krstev**. “A View on the Representation of Women in Serbian Newspaper Texts”. In Tamás Váradi, Judit Kuti and Max Silberztein (eds.) *Application of Finite-State Language Processing: Selected Papers from the 2008 International Nooj Conference*, pp. 160-170, Cambridge Scholars Publishing, 2010. ISBN (13) 978-1-4438-2573-3.
 - 2.34 **Cvetana Krstev**, Milan Vasiljević, “Statistika i evaluacija u bibliotekama” in *INFOteka : časopis za informatiku i bibliotekarstvo*, vol. XI, No. 2, pp. 55-67, Beograd, 2010.
 - 2.35 **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.
 - 2.36 Ranka Stanković, Ivan Obradović, **Cvetana Krstev**, Duško Vitas, “Production of morphological dictionaries of multi-word units using a multipurpose tool”, In: *Proceedings of the Computational Linguistics-Applications Conference*, October 17–19, 2011. Jachranka, Poland, eds. K. Jassem, P. W. Fuglewicz, M. Piasecki and A. Przepiórkowski, pp. 77 – 84, Polish Information Processing Society, ISBN 978-83-60810-47-7, 2011.
 - 2.37 **Cvetana Krstev**, Duško Vitas, Ivan Obradović, Miloš Utvić, “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, FSMNLP 2011*, Blois, France, July 12-15, 2011. eds. Andreas Maletti and Matthieu Constant, Association for Computational Linguistics, ISBN 978-3-642-14769-2, pp. 48-56, 2011.
 - 2.38 **Cvetana Krstev**, Duško Vitas, Aleksandra Trtovac, “Orwell’s 1984 – the Case of Serbian Revisited”, in *Proceedings of 5th Language & Technology Conference*, November 25-27, 2011, Poznań, Poland, ed. Zygmunt Vetulani, ISBN 978-83-932640-1-8, pp. 570-574, Fundacja Uniwersytetu im. A. Mickiewicza, Poznań, 2011.
 - 2.39 Ranka Stanković, Miloš Utvić, Duško Vitas, **Cvetana Krstev**, and Ivan Obradović. “On the Compatibility of Lexical Resources for Nooj”. In Kristina Vučković, Božo Bekavac and Max Silberztein (eds.) *Automatic Processing of Various Levels of Linguistic Phenomena: Selected Papers from the 2011 International Nooj Conference*, pp. 96-108, Cambridge Scholars Publishing, 2012. ISBN (13) 978-1-4438-3711-8.
 - 2.40 Ranka Stanković, **Cvetana Krstev**, Ivan Obradović, Aleksandra Trtovac and Miloš Utvić, “A Tool for Enhanced Search of Multilingual Digital Libraries of E-journals”, *Proceedings of the 8th International Conference on Language Resources and Evaluation, LREC 2012*, Istanbul, Turkez, 23--25 May 2012, 2012, eds. Nicoletta Calzolari *et al.*, ISBN 978-2-9517408-7-7.
 - 2.41 Maciej Ogrodniczuk, Radovan Garabík, Svetla Koeva, **Cvetana Krstev**, Piotr Pęzik, Tibor Pintér, Adam Przepiórkowski, György Szaszák, Marko Tadić, Tamás Váradi,

- Duško Vitas. "Central and South-European Language Resources in META-SHARE". in *INFOtheca : Journal of Librarianship and Informatics*, vol. XIII, No. 1, pp. 3-26, Belgrade, 2012. ISSN 1450-9687
- 2.42 **Cvetana Krstev**, Jelena Jaćimović and Duško Vitas, "Recognition and normalization of some classes of named entities in Serbian", *BCI '12 Proceedings of the Fifth Balkan Conference in Informatics*, ACM New York, NY, USA, 2012, pp. 52-57, eds. Mirjana Ivanović and Zoran Budimac, ISBN 978-1-4503-1240-0, DOI 10.1145/2371316.2371327
- 2.43 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
- 2.44 Duško Vitas, **Cvetana Krstev**, "Derivational Morphology in E-Dictionaries of Serbian", In: *Proceedings of the 32nd International Conference on Lexis and Grammar, September 10-14, 2013, Faro, Portugal*, eds. Jorge Baptista and Mario Monteleone, pp. 177-184, Universidade do Algarve, Faro, Portugal, 2013.
- 2.45 **Cvetana Krstev**, Ivan Obradović, Miloš Utvić, Duško Vitas, "A system for named entity recognition based on local grammars", *J Logic Computation* first published online February 19, 2013, Oxford Journals, doi:10.1093/logcom/exs079

(2.г) Наставни материјал

1. Наставни материјал за предмете Информатика за библиотекаре 1 и 2 доступан на вебу (теме: Бројеви и бројни системи; Основни појмови из теорије скупова; Операције и релације; Релације, функције и бесконачност; Исказни рачуна);
2. Наставни материјал за предмете Информатички практикум 1 и 2 доступан на вебу (теме: Историјат настанка рачунара; Рачунар; Софтвер; Мреже);
3. Наставни материјал за предмет Дигитални текст 1 доступан на вебу (теме: Кратак увод у типографију; Шта треба знати о LaTeX-у; О кодирању симбола писања);
4. Наставни материјал за предмет Дигитални текст 2 доступан на вебу (теме: Интернет; Машине за претраживање; CSS – каскадни стилски листови);
5. Наставни материјал за предмет Структура информација 1 доступан на вебу (теме: Кодирање бројева за потребе рачунања; Алгоритми; Контролне структуре; Низови и листе; Дрвета);
6. Наставни материјал за предмет Информатички практикум 4 доступан на вебу (теме: О обележавању текста, Увод у XML);
7. Наставни материјал за предмет Базе података и библиотечки информациони системи доступан на вебу (теме: Увод у предикатски рачун; Увод у базе података; Модел ентитет-везе; Релациони модел);
8. Наставни материјал за предмет Проналажење информација доступан на вебу.

(2.д) Учешће на научним пројектима од 2008. године

1. 2006-2010: Научно-истраживачки пројекат 148921А "Теоријско-методолошки оквир за модернизацију описа српског језика", учесници: Математички факултет, Филолошки факултет, Рударско-геолошки факултет (Београд) (категорија истраживача: А1)
2. 2011-2014: Научно-истраживачки пројекат ИИИ-47003 "Инфраструктура за електронски подржано учење" (учесници: из београда: ФОН, Математички факултет, Филолошки факултет, Рударско-геолошки факултет, Технолошки факултет, из Новог Сада: ПМФ и Факултет техничких наука, из Ниша: Електронски факултет, Технички факултет у Чачку)

3. 2011-2014: Научно-истраживачки пројекат ОН178006 “Српски језик и његови ресурси: теорија, опис и примене” учесници: Математички факултет, Филолошки факултет, Рударско-геолошки факултет (Београд) (категорија истраживача: А1)
4. 2010-2011: Пројекат у оквиру програма “Павле Савић” билатералне сарадње са Француском “Паралелни француско-српски корпус” (у сарадњи са Универзитетом у Арасу)
5. 2011-2013: CESAR – Central and South-East European Resources (ICT Policy Support Programme, Grant agreement no.: 271022)
6. 2013-2017: PARSEME - COST Action IC1207 “PARSEME: PARSing and Multi-word Expressions. Towards linguistic precision and computational efficiency in natural language processing”, Management Committee member, Stirling Committee member, 2013-2017

(2.ђ) Учешће у комисијама за одбрану магистарских и докторских радова од 2008. године

Докторске тезе – члан комисије

1. Ранка Станковић: *Модели експанзије упита над текстуелним ресурсима*, Математички факултет, 2009.
2. Александра Павловић: *Међубиблиотечка позајмица у Србији: развој и трансформација у складу са трансформацијом комуникационог система у науци*, Филолошки факултет, 2009.
3. Гордана Рудић: *Формална спецификација ограничења у моделима података библиографских записа*, Филолошки факултет, 2010.

Магистарски радови - ментор

1. Сања Антонић: *Развој информатичке семантичке мреже за област биомедицине*, Филолошки факултет, 2008.
2. Сандра Гуцул-Милојевић: *Препознавање једне класе именованих ентитета у електронским текстовима на српском језику*, Филолошки факултет, 2008.
3. Младен Божовић: *Рачунарско-лингвистичке методе поравнавања паралелних текстова и њихова примена на енглеско-српски језички пар*, Филолошки факултет, 2010.
4. Зоран Ристовић: *Примена паралелизованих текстова у настави енглеског језика у основној школи*, Филолошки факултет, 2012.

Магистарски радови – члан комисије

1. Јелена Бајић: *Morphological Aspects of Netspeak*, 2009.
2. Марина Ђукић: *Могућности и ограничења ауторских програма у настави немачког језика као страног помоћу рачунара*, 2010.
3. Карла Селихар: *Библиографија педагошког часописа 'Школски лист' 1858-1910*, 2011.
4. Ивана Спичановић: *Референсни рад у библиотекама*, 2013.

Мастер радови

1. Љиљана Мацура: *Електронски извори у научној читаоници Народне библиотеке Србије*, 2009;

2. Биљана Калезић: *Софтверска пиратерија у Србији*, 2009;
3. Адам Софронијевић: *Неки елементи бизнис-модела на примеру Универзитетске библиотеке*, 2009;
4. Предраг Ђукић: *Веб 2.0 и библиотеке*, 2010;
5. Душица Рајчевић: *Пројекат Светске дигиталне библиотеке — идеја, значај, визија*, 2010;
6. Олга Николић, *Интернационална дечија дигитална библиотека*, 2010;
7. Јелена Ничић: *Каталогизација дигиталних ресурса*, 2010;
8. Милан Васиљевић: *Набавка и организовање електронских извора у библиотекама — светска пракса*, 2010;
9. Миљана Ђорђевић: *Нормативна контрола у библиотечко-информационим системима*, 2010;
10. Марија Стиковић: *Циклус формирања дигиталних збирки*, 2010;
11. Милица Васић: *Утицај друштвених мрежа интернета на друштво*, 2010;
12. Јелена Марковић: *Пројекат "Гугл Књиге", Каталогизација дигиталних ресурса*, 2010;
13. Јелена Петровић: *Пројекат "Гутенберг"*, 2010;
14. Ружица Петровић: *Институционални репозиторијуми*, 2011;
15. Ивана Ковачевић: *Центар за дигитализацију Југословенске кинотеке*, 2011;
16. Јелена Андоновски: *Еуропа — врата до европског културног наслеђа*, 2011;
17. Сандра Дубравчић: *Застаревање дигиталних података*, 2011;
18. Мина Милић: *Принципи и технике функционисања веб претраживача уз осврт на Google, Yahoo и Bing*, 2011;
19. Биљана Ђорђевић: *Системи за управљање садржајем*, 2011;
20. Зорица Јанковић: *Корисничка сучеља електронских каталога с посебним освртом на Специјалну библиотеку Музеја града Београда*, 2012;
21. Биљана Лазић: *Архиви медијских кућа у свету*, 2012;
22. Дамир Симончић: *Књижевни часописи у дигиталним библиотекама Србије; актуелна пракса и нове могућности*, 2012.

Докторске тезе које су тренутно у изради под менторством др Цветане Крстев

1. мр Сања Антонић: *Језички алати за проналажење информација из области биомедицине на српском језику*, тема одобрена октобра 2008.
2. мр Сандра Гуцул-Милојевић: *Локалне граматике у екстракцији информација*, тема одобрена март 2009.
3. мр Милош Утвић: *Изградња референтног корпуса савременог српског језика*, тема одобрена октобра 2009.
4. мр Зоран Ристовић: *Кумулативни ефекти експлоатације вишејезичних корпуса у настави страних језика*, тема одобрена априла 2013.
5. мр Адам Софронијевић: *Нова парадигма сарадње у библиотекама*, тема одобрена априла 2013.
6. мр Небојша Васиљевић: *Аутоматска обрада правних текстова на српском језику*, тема одобрена јуна 2013.

(2.е) Софтверски алати и ресурси

- 2.е.1 Српски морфолошки електронски речник простих речи у Multext-East формату;
- 2.е.2 Опис морфологије именских вишечланих речи и израда електронског речника вишечланих речи у DELAC формату (речник тренутно има 12,000 одредница);

2.e.3 Корпус савременог српског језика доступан преко веба од 118 милиона текућих речи (коаутор);

2.e.4 Вишејезички корпуси (коаутор):

- *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;
- *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;
- Паралелизовани текст Верновог романа “Пут око света за 80 дана” на 18 језика.

2.e.5 Морфолошки обележени и лематизовани текстови:

- Џорџ Орвел *1984*;
- Жил Верн „Пут око света за 80 дана“.

2.e.6 Алати (коаутор):

- *Bibliša* - претраживање поравнатих текстова из електронских часописа (<http://hlt.rgf.bg.ac.rs/Biblisha/>);
- *VebRanka* - претраживање веба проширивањем упита (<http://hlt.rgf.bg.ac.rs/VeBranka/>);
- *NERosetta* - претраживање паралелних текстова обележених именованим ентитетима (<http://www.korpus.matf.bg.ac.rs/nerosetta/>);
- *NERanka* - обележавање корисникових текстова именованим ентитетима (<http://hlt.rgf.bg.ac.rs/VeBranka/NERanka.aspx>).

(2.ж) Рецензије радова

Кандидат је, као члан програмског одбора рецензирала од 3-5 радова по конференцији на следећим научним скуповима:

- 6th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics December 7-9, 2013, Poznań, Poland
- 32th International Conference on Lexis and Grammar, September 10-14, 2013, Faro, Portugal
- Eighth Language Technologies Conference, October 8 - 9, 2012, Ljubljana, Slovenia
- 31th International Conference on Lexis and Grammar, September 19-22, 2012, Nové Hrad, Czech Republic
- CLoBL 2012: Workshop on Computational Linguistics and Natural Language Processing of Balkan Languages, September 16, 2012, Novi Sad, Serbia
- QUALICO 2012 - Structures and Processes in Language and Text: Quantitative Linguistics Perspectives, April 26-29, 2012, Belgrade, Serbia
- 30th International Conference on Lexis and Grammar, October 5-8, 2011, Nicosia, Cyprus
- Seventh conference on Language Technologies, Ljubljana, October 2010
- 29th International Conference on Lexis and Grammar, September 15-18, 2010, Belgrade, Serbia
- Multiword Expressions: From Theory to Applications (MWE 2010), Workshop associated with the 23rd International Conference on Computational Linguistics (COLING 2010), August 23-27, 2010, Beijing, China
- Exploitation of multilingual resources and tools for Central and (South) Eastern European Languages, Workshop associated with the LREC conference, Malta, 17-23 May 2010
- 28th Lexis & Grammar Conference, Bergen, Norway, September 30 – October 2, 2009.

- Multilingual resources, technologies and evaluation for central and Eastern European languages, September 26, 2009, a workshop associated with the RANLP (Recent Advances in Languages Processing) Conference
- Sixth Language Technologies Conference IS- LTC, Ljubljana, Slovenia, October 16-17, 2008

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, чији је један од оснивача, а 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 Цветане Крстев 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, електронских речника и онтологија, затим из рачунарске морфологије и морфосинтаксе, парцијалне синтаксичке анализе (именовани ентитети) као и пројектовање ових резултата на осавремењивање библиотечких информатизованих делатности (нпр. дигиталне библиотеке).

1. Прегледни радови. Рад под бројем 2.8 је монографија објављена са групом коаутора у оквиру европског пројекта Мета-нет, а са циљем да прикаже ниво информатичке описаности српског у односу на потребе информационог друштва. Студија прегледно приказује у посебним поглављима опасности и изазове које доносе језичке технологије, статус српског језика у информационом друштву и језичке технологије за српски језик. Студија се закључује упоредном анализом стања у 31 европском језику, која је дело колективног рада у оквиру пројекта Мета-нет. Ова студија, као и остале студије из ове серије су биле основа за утврђивање европске стратешке агенде за вишејезично друштво до 2020. године.

У раду под бројем 2.41 приказани су језички ресурси који су изграђени у оквиру европског пројекта CESAR и постављени на вебу (под СС-лиценцом) кроз систем МЕТА-размена.

2. Корпуси. Кандидат је учествовао или самостално изградио више корпуса. С једне стране, ту је *Корпус савременог српског језика* који је доступан преко веба са сервера смештеног на Математичком факултету, као и више специјализованих корпуса од којих је посебно занимљив корпус изабраних текстова из дневника Политика у периоду 1904-2010. Осим корпуса који је доступан преко Интернета, кандидат је, користећи систем Unitex, развио и друкчији поглед на корпус заснован на експлоатацији информација у електронском морфолошком речнику. Ова два аспекта експлоатације корпуса описана су у раду под бројем 2.43. Овде је посебно илустрована предност коју доноси анализа корпуса коришћењем електронских речника и локалних граматика над уобичајеним системима за анализу заснованим на статистичким методама.

Други групу радова чине радови посвећени изградњи и експлоатацији паралелних корпуса. Кандидат је учествовао у изградњи више паралелних корпуса и колекција паралелизованих текстова. У раду под бројем 2.1 је дат преглед састављених паралелних корпуса коришћењем система Xalign, као и могућности његове надградње. Овај рад је био полазиште групи млађих сарадника да развије систем ACIDE који поједностављује процес контроле паралелизације, а паралелизовани текст представља у TMX- и html-форматима поред TEI-формата. Рад 2.7, објављен у монографији посвећеној Валтеру фон Хану, дат је преглед паралалелизованих корпуса у којима учествује српски као један од језика у битексту (енглеско-српски, француско-српски, хрватско-српски, српско-српски, пољско-српски), као и преглед метода поравнавања и начина експлоатације (кроз системе IMS CQP и Unitex). У раду под бројем 2.35 дат је приказ паралелизованог енглеско-српског корпуса литерарних текстова и примери његове експлоатације коришћењем система Unitex. У раду под бројем 2.17, полазећи од паралелизоване верзије романа Жила Верна "Пут око света за 80 дана" анализира се систем морфолошке анотације текстова у различитим језицима (француски, енглески, српски, бугарски, грчки, пољски) за које постоје електронски морфолошки речници. Ова анализа указује на одсуство нормализоване нотације како на нивоу обележавања врста речи, тако и у обележавању морфолошких, граматичких и семантичких својстава у електронским речницима различитих језика што отежава развој система за упаривање на нивоу речи и евентуални трансфер локалних граматика из једног језика у други. Рад 2.38 полази од система морфолошке анотације који је развијен у оквиру пројекта Multext-East, а у коме је предложен један позициони систем етикетирања граматичких категорија за језике Источне и Централне Европе на нивоу простих речи (тј. непрекинутих секвенци алфабетских симбола). Предложена је ревизија овог стандарда која би водила рачуна о анотирању полилексичких јединица и њихових својстава, а у складу са текућим процесом међународне стандардизације. Радови 2.19 и 2.29 приказују језичке ресурсе - паралелизоване и упоредне корпусе, као и језичке моделе за балканске језике - са циљем да се изграде системи за њихово непосредно аутоматско превођење (без интерлингвалног посредовања енглеског језика). У раду 2.9 је, с обзиром на искуства у обради српског, посебно у обради текстова из 19. века, приказана методологија која се може користити у обради јудео-шпанског (транслитерација, нормализација, коришћење постојећег електронског речника шпанског, начина његове надградње, и сл). У раду под бројем 2.40 приказан је веб-апликација *Bibliša* која омогућава да се на вебу постављају проширени упити над паралелизованом верзијом чланака објављених у часопису *Инфотека*, чији је уредник др Цветана Крстев. 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 и да на тај начин обезбеди повољнију прецизност у претраживању.

3. 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 развио систем локалних граматика и каскада трансдуктора који омогућава да се за српски језик препознају широке класе именованих ентитета. У раду 2.20, кандидат је дао прву апроксимацију препознавања појединих класа именованих ентитета користећи локалне граматике развијене под системом NooJ, а у раду под бројем 2.33 је дата примена реализованог система на препознавање појављивања женских имена у корпусу дневног листа *Политика*. Проширивање система за препознавање именованих ентитета је обухватило и аспкете њихове

вишејезичне употребе, што је изложено у раду под бројем 2.3. Ту је приказана лексичка база *Prolex* која је развијена у сардњи са Универзитетом Франсоа Рабле из Тура. Основна идеја у конструкцији базе је раздвајање језичког нивоа на коме се реализују властита имена у одређеном језику од концептуалног нивоа који је заједнички за више језика. Пренос концептуалних информација о једном властитом имену се преноси из једног језика у други коришћењем јединственог идентификационог броја властитог имена. У раду се анализирају и различите релације које повезују властита имена. Проширење система за препознавање именованих ентитета временски изразима (датум и период) и то у вишејезичној анализи текста је дато у раду 2.28. Овде су, користећи се паралелизованим француско-српским корпусом, анализирани именовани ентитети који изражавају одређени временски период у два језика, а са циљем да се успоставе механизми за евентуално аутоматско превођење ових величина из једног језика у други. У раду под бројем 2.42 размотрен је проблем стандардизованог приказивања препознатих именованих ентитета ослањајући се на одговарајуће норме Међународне организације за стандардизацију (ISO). Рад приказује систем коначних трансдуктора који за дати улазни текст генерише xml-етикете чији су атрибути стандардизоване вредности препознатих ентитета. Овај корак у препознавању именованих ентитета је за друге језике углавном још у повоју. Добијени резултат омогућава да се екстраховани ентитет похрани на начин који омогућава високо ефикасну екстракцију информација. Радови под бројем 2.37 и 2.45 разматрају из теоријског угла улогу локалних граматика и каскада коначних трансдуктора у препознавању именованих ентитета.

Поред радова у којима се описују методе препознавања именованих ентитета, кандидат је са сарадницим развио више веб-апликација које илуструју постигнуте резултате. Посебно се истичу апликација *NERanka* у којој корисник може да пошаље веб-сервису свој текст на српском, а да му се као резултат врати текст са обележеним ентитетима. Друга значајна апликација је *NERosetta* која допушта поређење резултата добијених у препознавању именованих ентитета у различитим језицима полазећи од паралелизованих текстова.

4. Семантичке мреже. Поред већ споменуте лексичке мреже властитих имена *Prolex*, кандидат др Цветана Крстев је развила значајан део семантичке мреже *WordNet*. Ова мрежа, преко јединственог идентификационог броја нумерише концепте и њихову лексикализацију у одређеном језику. Како је мрежа настала првобитно за енглески језик, основни идентификатори концепата се преузимају из енглеског. Сваком концепту одговара скуп квази-синонима (који се назива **синсет**), а синсетови су повезани различитим семантичким релацијама. Српски део *WordNet*-а је настао кроз појекат Балканет Европске комисије и по завршетку пројекта је имао око 7.000 синсетова, док их данас има преко 20.000. Конструкција ове базе за српски је вршена ручно у чему је учествовао велики број волонтера чији рад је контролисао главни уредник *WordNet*-а, др Цветана Крстев (приказано у раду 2.14). Квантитативна анализа садржаја српског дела ове лексичке базе је дата у раду под бројем 2.5. У раду 2.13 је размотрен проблем додавања синсетова за концепте који немају своју лексикализацију у принстонском *WordNet*-у, а имају је у језицима Балкана. То су нарочито концепти везани за флору и фауну Балкана, храну и родбинске везе. Проширење овог истраживања је дато у раду под бројем 2.26. У раду 2.23 је описано проширење српског *WordNet*-а биомедицинском терминологијом. У раду 2.15 је размотрена веза између система морфолошких електронских речника и *WordNet*-а како би се на аутоматски начин одредницама из електронског речника приписала семантичка својства описана кроз *WordNet* и, обратно, генерисале деривационе релације у *WordNet*-у на основу описа одредница у електронском речнику. Полазећи од оваквог обогаћивања лексичких ресурса, изграђен је систем за експанзију упита који се поставља претраживачким машинама. Овај

систем, назван *Вебранка*, омогућава да се за полазни упит изврши његово морфолошко (сви облици флективне парадигме) и семантичко проширење, да се одреди евентуални енглески еквивалент упита, и да се тако проширени упит упуту претраживачкој машини (предефинисана верзија претраживачке машине је *Гул*). Пример примене овог система у развоју Система научних и технолошких информација Србије је дат у раду 2.25 на примерима непотпуних одговора из каталога Народне библиотеке.

5. 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 *бела врана*, која се у речнику простих речи јавља као низ од две засебне одреднице - *бео*, придев обележен као боја, и *врана*, именица која означава зоолошки појам, у речнику полилексичких јединица јавити као једна одредница и то као именица која означава особу. Слично, израз *с времена на време* 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 *бела врана*, 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 Цветана Крстев је у раду под бројем 2.4 описала принципе конструкције речника полилексичких јединица у српском користећи програмски систем *Multiflex* у чијем је развоју учествовала, а који се темељи на механизму унификације. Развој речника полилексичких јединица и начин одређивања њихових флективних класа је изложен у радовима 2.10 и 2.27. У раду под бројем 2.16 се анализира систем анотације који укључује речник полилексичких јединица, а који представља значајан напредак у обликовању излаза из морфолошког анализатора који је потенцијални улаз за синтаксички анализатор. Интеграција механизма генерисања полилексичких јединица у лексикографску станицу *LeXimir* приказана је у раду 2.36. Кандидат је развио посебан подсистем који, у оквиру система *LeXimir*, 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 2.31 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.

6. Морфологија. Кандидат др Цветана Крстев је изградила електронски морфолошки речник српског који је намењен аутоматској анализи текста. Овај речник садржи око 130.000 одредница и преко 4.500.000 флективних облика од којих је сваки пропраћен одговарајућом граматичком информацијом. Речник је састављен у такозваном LADL-формату који се користи у одређеним програмским системима за анализу текста (*Intex*, *NooJ*, *Unitex*). У раду под бројем 2.30 размотрен је систем декларација обележја (FSD), сагласан са препорукама ТЕИ-конзорцијума, са потребама конверзије електронског речника из специфичног LADL -формата у стандардизоване облике и дат је опис морфолошких обележја српског језика у овом формализму. Ово истраживање је омогућило конверзију електронског речника у стандардизовани LMF-

формат који је препоручен стандардима ISO-a и основа је за размену и преносивост овакве врсте података. У раду под бројем 2.39 се анализирају недоследности у кодирању граматичких својстава у имплементацији електронских речника под системом *NooJ*, а у раду под бројем 2.16 се разматра општи модел аутоматске морфосинтаксичке анотације засноване на садржају електронских речника. Насупрот статистичким обележивачима, анотација заснована на речнику пружа могућност врло прецизног морфосинтаксичког описа улазног текста. Рад разматра предложени модел на примерима из српског и бугарског језика. У раду под бројем 2.44 је описан систем морфолошких граматика у облику коначних трансдуктора којима се апроксимирају феномени регуларне деривације у српском језику.

7. Библиотечки системи. У раду 2.2, кандидат др Цветана Крстев разматра могућност употребе грађе из дигиталних библиотека у Србији у лингвистичким истраживањима, а у раду под бројем 2.22 указује на проблеме које постављају нетранспарентни формати дигиталних текстова за њихову даљу употребу. У раду 2.18 кандидат указује на предност састављања библиографских података у информатизованом облику. У раду 2.34 разматрају се статистичке обраде и методе евалуације у библиотекама. Рад 2.32 излаже предности употребе слободног над власничким софтвером са посебним освртом на стање у домаћим библиотекама.

8. Лексикографија. Поред наведених објављених радова, кандидат је идејни аутор програмског система *LeXimir* који интегрише различите функције за рад са лексичким ресурсима и њихову примену у обради корпуса. Овај систем представља прототип лексикографске радне станице који допушта да се хетерогени лексички ресурси интегришу, примене и провере кроз обраду једнојезичног или паралелизованог корпуса.

Цитираност радова

Радови др Цветана Крстев су цитирани према базама података *Science Citation Index* и *Web of Science* укупно 47 пута без аутоцитата до октобра 2013. године. Број цитата изван ових база, а према <http://scholar.google.com> је значајно већи.

Научна и стручна активност. На основу података из биографије др Цветана Крстев, овде издвајамо следеће њене научне и стручне активности. Од 2007. године, др Цветана Крстев уређује као главни и одговорни уредник часопис *Инфотека* Заједнице универзитетских библиотека Србије, а који излази на српском и енглеском језику са значајним учешћем реномираних иностраних аутора. Члан је Комисије за корпус Одбора за стандардизацију српског језика као један од твораца Корпуса савременог српског језика и више паралелизованих корпуса у којима учествује српски језик. Од 1981. године учествовала је у 12 научних и развојних домаћих и 10 међународних научних пројеката. Од 2013. до 2017. је члан Надзорног одбора COST акције IC1207 PARSEME. Учествовала је на великом броју међународних и домаћих конференција, а на двадесет међународних конференција је била члан програмског одбора. О научним резултатима говори и цитираност на основу база података *Science Citation Index* и *Web of Science*, према којима др Цветана Крстев има 47 цитата без аутоцитата. Непосредно је организовала четири међународне конференције у Београду и Новом Саду. Држала је предавања по позиву на више европских универзитета, а као гостујући професор боравила на Универзитету Париз-Исток и у Институту за језик Бугарске академије наука. Поред овога, др Цветана Крстев активно учествује у развоју алата и ресурса намењених обради и истраживању српског језика.

Наставна и наставничка активност. Цветана Крстев предаје тринаест предмета на основним студијама и то *Информатика за библиотекарe I и II, Информатички практикум I-IV, Дигитални текст I и II, Структуре информација I и II, Базе података и библиотечки информациони системи, Проналажење информација и Мултимедијални документи*, а такође је укључена у наставу на мастер и докторским студијама. За потребе студената припремила је обиман наставни материјал који им је доступан преко веба и који се сваке године допуњује и ажурира. Са студентима завршне године основних студија је у оквиру предмета *Мултимедијални документи* израдила четири мултимедијална документа: (1) *Александар Аџа Поповић*; (2) *Култне радио емисије*; (3) *Поп Ђира и поп Спира*; (4) *Како су нас засмејавали (о Лоли Ђукићу и Новаку Новаку)*. Под руководством др Цветане Крстев студенти су своје радове представили на конференцијама које сваке године организује Филолошки факултет, на семинарима за школске библиотекарe и у часопису *Инфотека* (са пратећим диском). О њеним наставничким активностима довољно говори чињеница да је под њеним менторством одбрањен једна докторска дисертација (коментор), седам магистарских и двадесет и два мастер рада. У овом тренутку под њеним менторством ради се шест докторских теза и већи број мастер радова. Учествовала је у више комисија за одбрану докторских теза и магистарских и мастер радова.

IV. Испуњеност других услова за избор

Увидом у приложену документацију, као и у рад и радове кандидата, Комисија закључује да су у потпуности испуњени формални (законски и статутарни) и суштински услови за избор др Цветана Крстев у звање редовног професора.

V. Предлог Комисије

На основу свега што је речено и што је Комисији познато, једногласно и свесрдно предлажемо Изборном већу Филолошког факултета Универзитета у Београду да др Цветану Крстев изабере у звање редовног професора за ужу научну област Библиотечка информатика, предмет Проналажење информација.

У Београду, 13. новембра 2013.

Комисија:

(др Александра Вранеш, ред. проф.)

(др Божо Ђорић, ред. проф.)

(др Гордана Павловић-Лажетић, ред. проф.)