



South Slavic Languages in the Digital Environment - JuDig
November 21-23, 2024, Belgrade, Serbia
CONFERENCE PROGRAMME

Јужнословенски језици у дигиталном окружењу - ЈуДиг
21 - 23. новембар 2024. Београд, Србија
ПРОГРАМ КОНФЕРЕНЦИЈЕ

Thursday, November 21st

08:30-09:30	Registration (1st floor aula)	
09:30-10:00	OPENING CEREMONY Conference Room (Sala za sednice)	
10:00-11:00	Invited talk: Ruslan Mitkov The evolution of Natural Language Processing: from rules through neural networks to generative AI. What does the future hold? Conference Room (Sala za sednice)	
11:00-11:30	<i>Coffee break</i>	
11:30-13:30	Digital corpora of South Slavic languages 1 Chair: Cvetana Krstev Room: 33	Language resources for South Slavic languages Chair: Katerina Zdravkova Conference Room (Sala za sednice)
11:30-11:50	Saša Moderc - Italijanski klitici i tagovanje: korisnička iskustva u radu s korpusom SERBITACOR3	Рада Стијовић, Ранка Станковић, Михаило Шкорић - Речник савременог српског језика: PCCJ
11:50-12:10	Душко Витас, Ранка Станковић, Цветана Крстев - Многа лица СрпКор-а	Jasmina Moskovljević Popović - Challenges in Developmental Data Annotation
12:10-12:30	Nevena Seković - Revision of Errors in the ITALSERB Learner Corpus	Милена Милинковић, Милица Иконић Нешић - Именовани ентитети у дигиталном корпусу просторних планова
12:30-12:50	Јелена Марковић - Ученички корпуси србофоних говорника енглеског језика у контрастивној анализи међујезика	Ranka Stanković, Jovana Radenović, Maja Ristić, Dragan Stankov - Creation of a Training Dataset for Question-Answering Models in Serbian
12:50-13:10	Kristina Ilić - Značaj paralelnih korpusa za istraživanje frazemskih konstrukcija u nemačkom i srpskom jeziku	Саша Марјановић, Дејан Стошић - Примена електронског конјугатора SERBOVERB у проучавању овладавања глаголском флексијом српскога језика
13:10-13:30	Milica Dinić Marinković, Milena Oparnica - DeciKo – Korpus knjiga za decu ranog i predškolskog uzrasta. Trenutno stanje i izazovi	Jelena Lazarević, Olivera Kitanović - Contrastive Analysis of Syntax Patterns in Comparable Football Corpora in Spanish and Serbian Languages

13:30-15:00	<i>Lunch break</i>	
15:00-16:00	Invited talk: Svetla Koeva LLM-based (Computational) Linguistics Research: Between Profit and Risk Conference Room (Sala za sednice)	
16:00-17:00	Language technologies for South Slavic languages 1 Chair: Jasmina Moskovljević Popović Room: 33	Workshop: Corpus analysis: textometry, TXM and other tools (part 1) by TESLA project team (University of Belgrade and JeRTeh) Conference Room (Sala za sednice)
16:00-16:20	Nikitas N. Karanikolas - Noun Phrase and Prepositional Phrase Chunking for the Greek Language with spaCy	
16:20-16:40	Jaka Čibej - First Steps Towards an Online Service for Automatic Morphological Inflection of Serbian and Croatian	
16:40-17:00	Ana Ostroški Anić, Ivana Brač - A Frame-Semantic Description of Cognition Verbs in Croatian	
17:00-17:20	<i>Coffee break</i>	
17:20-19:00	Grammar and lexicon of South Slavic languages in the context of NLP Chair: Ana Ostroški Anić Room: 33	Workshop: Corpus analysis: textometry, TXM and other tools (part 2) by TESLA project team (University of Belgrade and JeRTeh) Conference Room (Sala za sednice)
17:20-17:40	Марина Баги - Синтаксичка и семантичка анализа глагола оштетити и уништити из угла теорије семантике оквира	
17:40-18:00	Наташа Киш - Синтаксичко-семантичка анотација електронских корпуса српског језика	
18:00-18:20	Таня Нейчева - Творителен предикативен падеж в съвременния сръбски, руски и полски език (по данни от многоезичен онлайн речник)	
18:20-18:40	Svetlozara Leseva, Ivelina Stoyanova - Towards Classifying Activity Predicates Involving Change	
18:40-19:00	Maја Matijeвић - Od rječnika i korpusa do hiponimije i meronimije	
bronze sponsor	 INSTITUTE FOR STANDARDIZATION OF SERBIA	

Friday, November 22nd

09:30-10:30	Invited talk: Agata Savary Automatic identification of multiword expressions - recent progress and perspectives Conference Room (Sala za sednice)	
10:30-11:30	Artificial intelligence, language models and processing of South Slavic languages 1 Chair: Mihailo Škorić Room: 33	IT processing of South Slavic languages 1 Chair: Rada Stijović Conference Room (Sala za sednice)
10:30-10:50	Maram Alharbi, Ruslan Mitkov - Comparing Rule-based and Deep Learning Approaches for Understanding Sentiments	Anna Jouravel, Martin Meindl, Achim Rabus, Elena Renje - Creating a Multi-Layered Digital Framework for Analysing Church Slavonic Manuscripts
10:50-11:10	Danka Jokić, Ranka Stanković, Jelena Jaćimović - Knowledge Graphs in the Era of Large Language Models: Opportunities and Challenges	Владимир Поломац, Тамара Лутовац Казновац, Марко Милошевић, Ана Марија Павловић - Модели за аутоматско препознавање старе српске штампане и рукописне ћирилице: тренутно стање и будући задаци
11:10-11:30	Milica Ikonić Nešić, Miloš Utvić - Overview of the TESLA-NER-NEL-Gold Dataset: Showcase on Serbian-English Parallel Corpus	Katerina Zdravkova, Jana Kuzmanova - Sonority Based Syllabification of Macedonian and Serbian Language
11:30-11:50	<i>Coffee break</i>	
11:50-13:30	Artificial intelligence, language models and processing of South Slavic languages 2 Chair: Vladimir Polomac Room: 33	Digital corpora of South Slavic languages 2 Chair: Duško Vitas Conference Room (Sala za sednice)
11:50-12:10	Ana Kovačević - Lažne Vesti i Generativna Veštačka Inteligencija: rizici i moguća rešenja	Simon Krek, Carole Tiberius, Jaka Čibej, Ana Ostroški Anić, Ranka Stanković - Extension of the ELEXIS-WSD Parallel Sense-Annotated Corpus with South Slavic Languages: Challenges, Results, and Plans
12:10-12:30	Mihailo Škorić - New Language Models for South Slavic Languages	Dušanka Vujović, Branko Milosavljević - Korpus SRPKO kao tehnološka osnova za izradu rečnika savremenog srpskog jezika
12:30-12:50	Saša Petalinkar, Milica Ikonić Nešić - Automating Synset Example Creation: A Case Study with Serbian Wordnet and ChatGPT	Olja Perišić - Corpora for Learning Serbian as a Foreign Language in the Era of Artificial Intelligence
12:50-13:10	Milena Šošić, Ranka Stanković, Jelena Graovac - Social-Emo.SR: Emocionalna višeznačna kategorizacija konverzionih poruka sa društvenih mreža X i Reddit	Jelena Redli - Značaj digitalnog korpusa i jezičkih alata u jezičkoj analizi forenzičkih tekstova na srpskom jeziku
13:10-13:30	Никола Јанковић, Јована Иваниш - Употреба модела Whisper Large v3 Sr за транскрипцију говора на српском језику у програмском језику Пајтон на платформи Гугл колаб	Саша Марјановић, Дејан Стошић - Збирке текстова на српском језику у деветојезичном паралелном корпусу PARCOLAB

13:30-15:00	<i>Lunch break</i>	
15:00-16:00	Invited talk: Cvetana Krstev Odnos „nas“ i „drugih“ u korpusu SrpELTeC dobijen starim dobrim metodama Conference Room (Sala za sednice)	
16:00-17:00	IT processing of South Slavic languages 2 Chair: Olja Perišić Room: 33	Workshop: Named-entity recognition (NER) and linking with Wikidata (part 1) by TESLA project team (University of Belgrade and JeRTeh) Conference Room (Sala za sednice)
16:00-16:20	Saša Petalinkar - Automatic Synset Labeling for Serbian WordNet: Development, Preprocessing, and Feasibility Assessment	
16:20-16:40	Данило Алексић - Унапређене верзије програма „Ка минималним паровима“	
16:40-17:00		
16:00-17:20	<i>Coffee break</i>	
17:20-19:00	Language technologies for South Slavic languages 2 Chair: Danilo Aleksić Room: 33	Workshop: Named-entity recognition (NER) and linking with Wikidata (part 2) by TESLA project team (University of Belgrade and JeRTeh) Conference Room (Sala za sednice)
17:20-17:40	Marija Đokić Petrović, Mihailo St. Popović, Vladimir Polomac - Utilizing Named Entity Recognition for the Analysis of Serbian Archival Documents	
17:40-18:00	Martina Pavić - Frekvencija značenja pridjeva u hrvatskome medicinskom nazivlju – korpusno utemeljena analiza kolokacija	
18:00-18:20	Nikola Janković - Methodology for Creating Multilingual Parallel Corpus Based on Online Digital User Manuals: The Hilti User Manuals Corpus	
18:20-18:40	Marija Pantić - Veštački korpusi za mašinsku obuku alata za proveru gramatike	
18:40-19:00	Анђелка Зечевић, Анастасија Жунић, Кристина Милојевић - Стари текстови, нове технологије: дигитализација докумената на српском језику	
20:30-	<i>Conference dinner</i>	

Saturday, November 23th

10:00-11:20	Digital humanities Chair: Vasilije Milnović Room: 33	Workshop: Corpus Query Language (CQL): Lexical gaps in bilingual corpora (part 1) by Olja Perišić (University of Turin) Conference Room (Sala za sednice)
10:00-10:20	Милена М. Стојановић - Значења и положај појединих лексема из сфере вештачке интелигенције у лексичком фонду српског језика и перцепција нових технологија – изазови модерног времена	
10:20-10:40	Небојша Ратковић - Интеграција Википедије на српском језику у образовне системе и унапређење језичких технологија	
10:40-11:00	Ана Мihaljević - Hrvatski crkvenoslavenski jezik i glagoljica u digitalnome okružju	
11:00-11:20	Снежана Петровић, Мирјана Петровић-Савић, Ана Шпановић, Ленка Бајчетић, Матија Нешовић, Јована Тодорић - Дигитализација грађе Одбора за ономастику САНУ – значај, циљеви и први кораци	
11:20-11:40	<i>Coffee break</i>	
11:40-13:20	Digital humanities Chair: Aleksandra Marković Room: 33	Workshop: Corpus Query Language (CQL): Lexical gaps in bilingual corpora (part 2) by Olja Perišić (University of Turin) Conference Room (Sala za sednice)
11:40-12:00	Vasilije Milnović, Aleksandra Trtovac, Cvetana Krstev, Ranka Stanković, Duško Vitas - Improving Machine Understanding of Text and Finding Information in Historical Newspapers in Serbia	
12:00-12:20	Андреја Сагић - Установе културе у ери вештачке интелигенције	
12:20-12:40	Срђан Шућур, Јелена Марковић - Дигитализација српског књижевног наслеђа ијекавског изговора (1840–1920) при Центру за дигиталну хуманистику Филозофског факултета Пале (прва фаза)	
12:40-13:00	Miloš Košprdić, Gorana Gojić, Adela Ljajić, Dragiša Mišković - Razvoj modela semantičke pretrage za srpski jezik	
13:00-13:20	CONFERENCE CLOSING <i>Conference Room (Sala za sednice)</i>	
13:20-14:30	<i>Lunch break</i>	
14:30-16:00	Workshop: Creating Lexical Networks using Large Language Models (LLM) with a Focus on Synonym extraction (part 1) by Benedikt Perak (University of Rijeka), Dragana Špica (University of Pula), Conference Room (Sala za sednice)	
16:00-16:15	<i>Coffee break</i>	
16:15-17:30	Workshop: Creating Lexical Networks using Large Language Models (LLM) with a Focus on Synonym extraction (part 2) by Benedikt Perak (University of Rijeka), Dragana Špica (University of Pula), Conference Room (Sala za sednice)	