



South Slavic Languages in the Digital Environment – JuDig
October 26–28, 2026, Belgrade, Serbia
CONFERENCE PROGRAMME

Јужнословенски језици у дигиталном окружењу – ЈуДиг
26–28. октобар 2026, Београд, Србија
ПРОГРАМ КОНФЕРЕНЦИЈЕ

	Sunday, October 25th	
10.00–11.45	Workshop: CLASSLA-express 3.0 workshops on using CLARIN.SI resources in Spoken and web corpora in the analysis of language variation by Tanja Samardžić and Kristina Ilić (Part 1, Room 151)	
11.45–12.15	Coffee break	
12.15–14.00	CLASSLA-express 3.0 workshops on using CLARIN.SI resources in Spoken and web corpora in the analysis of language variation (Part 2, Room 151)	
Monday, October 26th		
08.30–09.00	Registration	
09.00–10.00	OPENING CEREMONY AND SPONSOR INTRODUCTION AND PRESENTATION Room 167 (1st floor)	
10.00–11.00	Invited talk: Prof. Dr Tiago Timponi Torrent Curation is also something you need: on the role of Language Resources in AI Room 167 (1st floor)	
11.00–11.30	Coffee break	
11.30–13.30	Language resources and tools for South Slavic languages Chair: Svetlozara Leseva Room 167 (1st floor)	Digital corpora of South Slavic languages Chair: Marko Simonović Room 101 (Library)
11.30–11.50	Ivelina Stojanova, Svetlozara Leseva – Approaches to Automatic Annotation of Semantic Roles and Syntactic Relations (en)	Jelena Andonovski, Aleksandra Trtovac – Imenice i kolokacije pridev–imenica u srpskom delu paralelnog korpusa časopisa Infoteka: tekstometrijski pristup (sr)
11.50–12.10	Marina Bagi, Ely Matos, Aleksandra Marković – A FrameNet Analysis of Serbian Verbs Kidati and Cepati: Frames vs. Traditional Word Sense Disambiguation (en)	Jovana Bazić – Korpus javnog diskursa u italijanskom i srpskom jeziku: struktura, metodologija i istraživački potencijal (sr)
12.10–12.30	Saša Petalinkar, Nikola Vulović – A Reproducible Codex Skill Workflow for Serbian Word Sense Disambiguation with Serbian WordNet (en)	Nevena Ceković – ItalSerb govorni korpus L2 italijanskog i uporedni subkorpus L1 srpskog (sr)
12.30–12.50	Coffee break	
12.50–13.10	Ranka Stanković, Mihailo Škorić, Rada Stijović – Rečnik savremenog srpskog jezika: od digitalne leksikografske baze do otvorene jezičke infrastrukture (sr)	Jaka Čibej, Simon Krek, Carole Tiberius, Irina Lobzhanidze, Verginica Mititelu, Jelena Kallas, Kertu Saul, Kadri Muischnek, Karoliina Jõgi, Ranka Stanković, Cvetana Krstev, Milica Ikonić Nešić, Aleksandra Markovic, Ana Ostroški Anić, Bartłomiej Alberski, Vladimir Cvetkoski, Voula Giouli, Rusudan Makhachashvili, Olha Kanishcheva – ELEXIS-WSD 2.0: Expanding the Lexicon–Corpus Interface with New Languages and Annotation Layers (en)
13.10–13.30	Milica Dinić Marinković, Cvetana Krstev, Aleksandra Marković, Ranka Stanković – Repozitorijum simila savremenog srpskog jezika – izgradnja i evaluacija mašinski potpomognute anotacije (sr)	Mihailo Škorić, Cosimo Palma – Wiki Dumps to Training Corpora: South Slavic Case (en)
13.30–15.00	Lunch break	
15.00–16.00	Corpus-based linguistic analysis Chair: Rada Stijović Room 167 (1st floor)	Workshop Building Enriched Frame-Based Resources: The FrameNet Brasil Framework
15.00–15.20	Olja Perišić – Momo Kapor u digitalnom okruženju: Književno prevođenje i ChatGPT (sr)	by Tiago Timponi Torrent (UFJF Linguística – FrameNet Brasil) Ely Matos (Universidade Federal de Juiz de Fora) and Marcelo Viridiano (INCT-TILDIAR/CNPq (National Institute for Science and Technology on Responsible Artificial Intelligence for Computational Linguistics and Information Treatment and Dissemination)) (Part 1, Room 151)
15.20–15.40	Andela Babić, Marko Simonović, Predrag Kovacević – Deverbal Nominalisations in the Timok Speech: A Preliminary Corpus Study (en)	
15.40–16.00	Marijana Đukić, Slobodan Novokmet – Poredbeni frazeologizmi sa komponentom pas: korpusno zasnovana analiza (sr)	
16.00–16.20	Coffee break	

16.20–18.00	Language resources and tools for South Slavic languages Chair: Dragana Špica Room 167 (1st floor)	Workshop Building Enriched Frame-Based Resources: The FrameNet Brasil Framework (Part 2, Room 151)
16.20–16.40	Benedikt Perak, Dragana Špica – Adapting an Agentic Pipeline for Croatian Language Teaching: From the School Dictionary to CEFR-Based Conversational Practice (en)	
16.40–17.00	Mihailo Škorić, Nikola Janković, Tanja Samardžić – Comparative Intrinsic Evaluation of LLM Tokenizers on Serbian Text (en)	
17.00–17.20	Milica Jevtić, Tamara Vučenović, Miloš Stojanović – From Literary Memory to Platformed Culture: South Slavic Diaspora Identity and Algorithmic Marginalization in Canadian Multicultural Literature (en)	
17.20–17.40	Snežana Petrović, Vladimir Polomac, Ana Španović, Jovana Todorić, Matija Nešović, Lenka Bajčetić – Korisnički interfejs za unos onomastičkih podataka: iskustva, izazovi i unapređenje modela (sr)	
17.40–18.00	Veljko Miljanić, Jovan Turanjanin, Mihailo Škorić – Merenje kvaliteta automatskog preslovljavanja latinice na ćirilicu (sr)	

Tuesday, October 27th		
09.30–10.30	Invited talk: Prof. Dr Petya Osenova PressMint: Towards Interoperable Corpora of Historical Newspapers Room 167 (1st floor)	
10.30–11.50	Digital corpora of South Slavic languages (historical and multimodal) Chair: Katerina Zdravkova Room 167 (1st floor)	Workshop TESLA: From Text to Knowledge Graphs Named Entity Recognition and Linking to Wikidata, Relation Extraction, and Knowledge Graph Construction for Serbian by: Ranka Stanković and Milica Ikonić Nešić (Part 1, Room 151)
10.30–10.50	Predrag Kovačević, Marko Simonović, Marija Đokić Petrović – Hypothesis-Driven Corpus Querying at Scale: A Method and Its Application to -nje Nominalizations in BCMS (en)	
10.50–11.10	Maarten Janssen, Anna Jouravel, Piroska Lendvai – Combining Scholarly Edition and NLP: Historical Multilingual Parallel Text Collections on TEITOK (en)	
11.10–11.30	Danka Jokić, Srđan Šućur, Jelena Marković <i>ADMIRE 2 Serbian: Изградња српског мултимодалног скупа идиома</i> (sr)	
11.30–11.50	Jovana Radenović, Biljana Rujević, Aleksandra Tomašević, Olivera Kitanović – Od (slike) teksta ka grafu znanja: digitalizacija i semantičko obogaćivanje enciklopedije Sveznanje (sr)	
11.50–12.10	Coffee break	
12.10–13.30	Digital humanities Chair: Ivelina Stojanova Room 167 (1st floor)	Workshop TESLA: From Text to Knowledge Graphs Named Entity Recognition and Linking to Wikidata, Relation Extraction, and Knowledge Graph Construction for Serbian (Part 2, Room 151)
12.10–12.30	Jana Kuzmanova, Sonja Gievska, Katerina Zdravkova – Exploring In-Context Learning for LLM Machine Translation for Macedonian (en)	
12.30–12.50	Dragana Kunčer – Translation direction in the Gesta Regum Sclavorum: Quantitative and qualitative approach (en)	
12.50–13.10	Đurđa Perović – Sistemi za odgovaranje na pitanja u javnim pozivima za finansiranje naučnoistraživačkih projekata: tekstometrijska i NLP analiza institucionalnog diskursa (sr)	
13.10–13.30	Irena Srdanović, Benedikt Perak, Dragana Špica – Croatian Tour-Guiding Corpus for AI-Augmented Japanese Language and Culture Learning (en)	
13.30–15.00	Lunch break	
15.00–16.00	Sentiment Analysis & Text Mining Chair: Milica Dinić Marinković Room 167 (1st floor)	AI, language models and processing of South Slavic languages Chair: Aleksandra Marković Room 101 (Library)
15.00–15.20	Miloš Košprdić, Isidora Dekić, Tijana Radović – Towards Fine-Grained Annotation of Directed Hate and Abusive Speech in Serbian (en)	Svetla Peneva Koeva – Modelling Expression of Eventuality Types in Bulgarian with Large Language Models: Challenges and Opportunities (en)
15.20–15.40	Milena Šošić, Ranka Stanković – Hierarchical Recognition of Moral Foundations in Serbian Literary Corpus SrpELTeC (en)	Katerina Zdravkova – Linguistic Annotation of South Slavic Languages with Large Language Models (en)
15.40–16.00	Ranka Stanković, Milena Šošić – Lexicon-Based Exploration of Emotion Dynamics in TEI Novels: Case Study on SrpELTeC Corpus (en)	Milica Gašić, Songbo Hu, Anđelka Zečević, Jovana Randelović, Ranka Stanković, Mihailo Škorić – srWOZ: Towards Task-Oriented Dialog Modelling in Serbian (en)
16.00–16.20	Coffee break	
16.20-18.00	AI4LANG — Integrating AI for the Digital Transformation of Language Teaching: Benedikt Perak (University of Rijeka, Project Coordinator); Damir Medved (Syntagent, WP2 Lead); Dragana Špica (Juraj Dobrila University of Pula, WP3 Lead); Ranka Stanković (JeRTeh, WP4 Partner) Room 167 (1st floor)	
21.00	Conference dinner	

Wednesday, October 28th		
9.30–10.30	Invited talk: Prof. Dr Francesca Frontini CLARIN and the Emerging European Ecosystem for Language Data and Technologies: The Italian Perspective Room 167 (1st floor)	
10.30–11.30	Sentiment Analysis & Text Mining Chair: Nevena Ceković Room 167 (1st floor)	Workshop TESLA: language resources Use of language resources developed within the framework of the TESLA project, with a focus on the TESLA portal by Mihailo Škorić and Biljana Rujević (Part 1, Room 151)
10.30–10.50	Jelena Redli – Anotacija moralnih vrednosti u srpskom onlajn diskursu: izazovi i perspektive (sr)	
10.50–11.10	Danka Jokić, Milica Andrić – SentroSa: korpus za detekciju ironije i sarkazma u kratkim tekstovima na srpskom jeziku (sr)	
11.10–11.30	Таня Нейчева – Графичната организация на текста като аналитичен обект при изграждането на диахронен корпус (върху материал от един новобългарски дамаскин (bg)	
11.30–12.00	CONFERENCE CLOSING	
12.00–12.20	Coffee break	
12.20–13.30		Workshop TESLA: language resources Use of language resources developed within the framework of the TESLA project, with a focus on the TESLA portal (Part 2, Room 151)