

Slavic NLP 2023 Programme

Time Schedule - 6.5. 2023

9:00 - 9:10	Introduction
9:10 - 10:30	<p>Regular papers I</p> <p>9:10 - 9:30 - Resources and Few-shot Learners for In-context Learning in Slavic Languages Michal Štefánik, Marek Kadlčík, Piotr Gramacki and Petr Sojka</p> <p>9:30 - 9:50 - Information Extraction from Polish Radiology Reports Using Language Models Aleksander Obuchowski, Barbara Klaudel and Patryk Jasik</p> <p>9:50 - 10:10 - Dispersing the clouds of doubt: can cosine similarity of word embeddings help identify relation-level metaphors in Slovene? Mojca Brglez</p> <p>10:10 - 10:30 - Named Entity Recognition for Low-Resource Languages - Profiting from Language Families Sunna Torge, Andrei Politov, Christoph Lehmann, Bochra Saffar and Ziyang Tao</p>
10:30 - 11:15	Coffee break
11:15 - 12:45	<p>Regular papers II</p> <p>11:15 - 11:35 - Too Many Cooks Spoil the Model: Are Bilingual Models for Slovene Better than a Large Multilingual Model? Pranaydeep Singh, Aaron Maladry and Els Lefever</p> <p>11:35 - 11:55 - On Experiments of Detecting Persuasion Techniques in Polish and Russian Online News: Preliminary Study [Online] Nikolaos Nikolaidis, Nicolas Stefanovitch and Jakub Piskorski</p> <p>12:05 - 12:25 - Can BERT eat RuCoLA? Topological Data Analysis to Explain [Online] Irina Proskurina, Ekaterina Artemova and Irina Piontkovskaya</p> <p>12:25 - 12:45 - Automatic text simplification of Russian texts using control tokens Anna Dmitrieva</p>
12:45 - 14:15	Lunch break
14:15 - 15:15	<p>Invited Talk</p> <p>Together we are stronger: Collaborative development of language resources and technologies for South</p>

	<p>Slavic languages Nikola Ljubešić, Tanja Samardžić</p>
15:15 - 15:20	<p>Shared Task overview</p> <p>Slav-NER: the 4th Cross-lingual Challenge on Recognition, Normalization, Classification, and Linking of Named Entities across Slavic languages Roman Yangarber, Jakub Piskorski, Anna Dmitrieva, Michał Marcińczuk, Pavel Přibáň, Piotr Rybak and Josef Steinberger</p>
15:20 - 15:45	<p>Pitch presentations of short/shared task papers</p> <p>MAUPQA: Massive Automatically-created Polish Question Answering Dataset [Online] Piotr Rybak</p> <p>TrelBERT: A pre-trained encoder for Polish Twitter [Online] Wojciech Szmyd, Alicja Kotyla, Michał Zobniów, Piotr Falkiewicz, Jakub Bartczuk and Artur Zygańdo</p> <p>Croatian Film Review Dataset (Cro-FiReDa): A Sentiment Annotated Dataset of Film Reviews Gaurish Thakkar, Nives Mikelić Preradović and Marko Tadić</p> <p>Machine-translated texts from English to Polish show a potential for typological explanations in Source Language Identification Damiaan Reijnaers and Elize Herrewijnen</p> <p>Target Two Birds With One SToNe: Entity-Level Sentiment and Tone Analysis in Croatian News Headlines Ana Barić, Laura Majer, David Dukić, Marijana Grbeša-Zenzerović and Jan Snajder</p> <p>Is German secretly a Slavic language? What BERT probing can tell us about language groups Aleksandra Mysiak and Jacek Cyranka</p> <p>Analysis of Transfer Learning for Named Entity Recognition in South-Slavic Languages Nikola Ivačić, Thi Hong Hanh Tran, Boshko Koloski, Senja Pollak and Matthew Purver</p> <p>WikiGoldSK: Annotated Dataset, Baselines and Few-Shot Learning Experiments for Slovak Named Entity Recognition David Suba, Marek Suppa, Jozef Kubik, Endre Hamerlik and Martin Takac</p> <p>Measuring Gender Bias in West Slavic Language Models Sandra Martinková, Karolina Stanczak and Isabelle Augenstein</p> <p>Exploring the Use of Foundation Models for Named Entity Recognition and Lemmatization Tasks in Slavic Languages Gabriela Pařka and Artur Nowakowski</p> <p>Large Language Models for Multilingual Slavic Named Entity Linking Rinalds Vřksna, Inguna Skadiņa, Daiga Deksnė and Roberts Rozis</p>
15:45 - 16:30	<p>Poster session open & Coffee break</p>

16:30 - 17:00	Poster session - part II
17:00 - 18:00	Findings papers 17:00 - 17:20 - Going beyond research datasets: Novel intent discovery in the industry setting [Online] Aleksandra Chrabrowa, Tsimur Hadeliya, Dariusz Kajtoch, Robert Mroczkowski and Piotr Rybak 17:20 - 17:40 - MLASK: Multimodal Summarization of Video-based News Articles Mateusz Krubiński and Pavel Pecina Regular papers III 17:40 - 18:00 - Comparing domain-specific and domain-general BERT variants for inferred real-world knowledge through rare grammatical features in Serbian Sofia Lee and Jelke Bloem
18:00	End of the workshop