

10th Workshop on Slavic Natural Language Processing (SlavicNLP 2025)

Vienna, Austria
31 July 2025

ISBN: 979-8-3313-2404-9

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2025) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2025)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>Identifying Filled Pauses in Speech Across South and West Slavic Languages</i> Nikola Ljubešić, Ivan Porupski, Peter Rupnik and Taja Kuzman	1
<i>Few-Shot Prompting, Full-Scale Confusion: Evaluating Large Language Models for Humor Detection in Croatian Tweets</i> Petra Bago and Nikola Bakarić	9
<i>GigaEmbeddings — Efficient Russian Language Embedding Model</i> Egor Kolodin and Anastasia Ianina	17
<i>PL-Guard: Benchmarking Language Model Safety for Polish</i> Aleksandra Krasnodebska, Karolina Seweryn, Szymon Łukasik and Wojciech Kusa	25
<i>Dialects, Topic Models, and Border Effects: The Rusyn Case</i> Achim Rabus and Yves Scherrer	38
<i>Towards Open Foundation Language Model and Corpus for Macedonian: A Low-Resource Language</i> Stefan Krsteski, Borjan Sazdov, Matea Tashkovska, Branislav Gerazov and Hristijan Gjoreski	44
<i>Towards compact and efficient Slovak summarization models</i> Sebastian Petrik and Giang Nguyen	58
<i>Adapting Definition Modeling for New Languages: A Case Study on Belarusian</i> Daniela Kazakouskaya, Timothee Mickus and Janine Siewert	69
<i>Bridging the Gap with RedSQL: A Russian Text-to-SQL Benchmark for Domain-Specific Applications</i> Irina Brodskaya, Elena Tutubalina and Oleg Somov	76
<i>Can information theory unravel the subtext in a Chekhovian short story?</i> J. Nathanael Philipp, Olav Mueller-Reichau, Matthias Irmer, Michael Richter and Max Kölbl	84
<i>When the Dictionary Strikes Back: A Case Study on Slovak Migration Location Term Extraction and NER via Rule-Based vs. LLM Methods</i> Miroslav Blšták, Jaroslav Kopčan, Marek Suppa, Samuel Havran, Andrej Findor, Martin Takac and Marian Simko	91
<i>DIACU: A dataset for the DIACHronic analysis of Church Slavonic</i> Maria Cassese, Giovanni Puccetti, Marianna Napolitano and Andrea Esuli	101
<i>Characterizing Linguistic Shifts in Croatian News via Diachronic Word Embeddings</i> David Dukić, Ana Barić, Marko Čuljak, Josip Jukić and Martin Tutek	108
<i>What Makes You CLIC: Detection of Croatian Clickbait Headlines</i> Marija Anđelić, Dominik Šipek, Laura Majer and Jan Snajder	116
<i>Gender Representation Bias Analysis in LLM-Generated Czech and Slovenian Texts</i> Erik Derner and Kristina Batistič	124
<i>REPA: Russian Error Types Annotation for Evaluating Text Generation and Judgment Capabilities</i> Alexander Pugachev, Alena Fenogenova, Vladislav Mikhailov and Ekaterina Artemova	136
<i>Fine-Tuned Transformers for Detection and Classification of Persuasion Techniques in Slavic Languages</i> Ekaterina Loginova	151

<i>Rubic2: Ensemble Model for Russian Lemmatization</i>	
Ilia Afanasev, Anna Glazkova, Olga Lyashevskaya, Dmitry Morozov, Ivan Smal and Natalia Vlasova	157
<i>Gradient Flush at Slavic NLP 2025 Task: Leveraging Slavic BERT and Translation for Persuasion Techniques Classification</i>	
Sergey Senichev, Aleksandr Boriskin, Nikita Krayko and Daria Galimzianova	171
<i>Empowering Persuasion Detection in Slavic Texts through Two-Stage Generative Reasoning</i>	
Xin Zou, Chuhan Wang, Dailin Li, Yanan Wang, Jian Wang and Hongfei Lin	177
<i>Hierarchical Classification of Propaganda Techniques in Slavic Texts in Hyperbolic Space</i>	
Christopher Brückner and Pavel Pecina	183
<i>Team INSAntive at SlavicNLP-2025 Shared Task: Data Augmentation and Enhancement via Explanations for Persuasion Technique Classification</i>	
Yutong Wang, Diana Nurbakova and Sylvie Calabretto	190
<i>LLMs for Detection and Classification of Persuasion Techniques in Slavic Parliamentary Debates and Social Media Texts</i>	
Julia Jose and Rachel Greenstadt	202
<i>Fine-Tuned Transformer-Based Weighted Soft Voting Ensemble for Persuasion Technique Classification in Slavic Languages</i>	
Mahshar Yahan, Sakib Sarker and Mohammad Islam	217
<i>Robust Detection of Persuasion Techniques in Slavic Languages via Multitask Debiasing and Walking Embeddings</i>	
Ewelina Ksiezniak, Krzysztof Wecel and Marcin Sawinski	224
<i>Multilabel Classification of Persuasion Techniques with self-improving LLM agent: SlavicNLP 2025 Shared Task</i>	
Marcin Sawinski, Krzysztof Wecel and Ewelina Ksiezniak	231
<i>SlavicNLP 2025 Shared Task: Detection and Classification of Persuasion Techniques in Parliamentary Debates and Social Media</i>	
Jakub Piskorski, Dimitar Dimitrov, Filip Dobranić, Marina Ernst, Jacek Haneczok, Ivan Koychev, Nikola Ljubešić, Michal Marcinczuk, Arkadiusz Modzelewski, Ivo Moravski and Roman Yangarber	254