

International Conference on Recent Advances in Natural Language Processing (RANLP 2025)

Natural Language Processing in the
Generative AI Era

Varna, Bulgaria
8-10 September 2025

Volume 1 of 2

ISBN: 979-8-3313-3437-6

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. (2026)

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>Harnessing Open-Source LLMs for Tender Named Entity Recognition</i>	
Asim Abbas, Venelin Kovatchev, Mark Lee, Niloofar Shanavas and Mubashir Ali	1
<i>On the Limitations of Large Language Models (LLMs): False Attribution</i>	
Tosin Adewumi, Nudrat Habib, Lama Alkhaled and Elisa Barney	11
<i>Candidate Profile Summarization: A RAG Approach with Synthetic Data Generation for Tech Jobs</i>	
Anum Afzal, Ishwor Subedi and Florian Matthes	22
<i>PersianSciQA: A New Dataset for Bridging the Language Gap in Scientific Question Answering</i>	
Safoura Aghadavoud Jolfaei, Azadeh Mohebi and Zahra Hemmat	32
<i>Multilingual Pre-training Meets Supervised Neural Machine Translation: A Reproducible Evaluation on English–French and Finnish Translation</i>	
Benyamin Ahmadnia, Yeswanth Soma and Hossein Sarrafzadeh	38
<i>Advancing Clinical Translation in Nepali through Fine-Tuned Multilingual Models</i>	
Benyamin Ahmadnia, Sumaiya Shaikh, Bibek Poudel, Shazan Mohammed and Sahar Hooshmand	48
<i>Advancing Active Learning with Ensemble Strategies</i>	
Naif Alatrush, Sultan Alsarra, Afraa Alshammari, Luay Abdeljaber, Niamat Zawad, Latifur Khan, Patrick T. Brandt, Javier Osorio and Vito D’Orazio	57
<i>Evaluating Large Language Models on Sentiment Analysis in Arabic Dialects</i>	
Maram I. Alharbi, Saad Ezzini, Hansi Hettiarachchi, Tharindu Ranasinghe and Ruslan Mitkov .	67
<i>From Posts to Predictions: A User-Aware Framework for Faithful and Transparent Detection of Mental Health Risks on Social Media</i>	
Hessam Amini and Leila Kosseim	75
<i>Beyond Methods and Datasets Entities: Introducing SH-NER for Hardware and Software Entity Recognition in Scientific Text</i>	
Aftab Anjum, Nimra Maqbool and Ralf Krestel	85
<i>Toponym Resolution: Will Prompt Engineering Change Expectations?</i>	
Isuri Anuradha, Deshan Koshala Sumanathilaka, Ruslan Mitkov and Paul Rayson	95
<i>HoloBERT: Pre-Trained Transformer Model for Historical Narratives</i>	
Isuri Anuradha, Le An Ha and Ruslan Mitkov	105
<i>A Framework for Fine-Tuning LLMs Using Heterogeneous Feedback</i>	
Ryan Aponte, Ryan A. Rossi, Shunan Guo, Franck Dernoncourt, Tong Yu, Xiang Chen, Subrata Mitra and Nedin Lipka	111
<i>Chakoshi: A Customizable Guardrail for LLMs with a Focus on Japanese-Language Moderation</i>	
Kazuhiro Arai, Ryota Matsui, Kenji Miyama, Yudai Yamamoto, Ren Shibamiya, Kaito Sugimoto and Yoshimasa Iwase	118
<i>KoWit-24: A Richly Annotated Dataset of Wordplay in News Headlines</i>	
Alexander Baranov, Anna Palatkina, Yulia Makovka and Pavel Braslavski	125

<i>Improving Estonian Text Simplification through Pretrained Language Models and Custom Datasets</i> Eduard Barbu, Meeri-Ly Muru and Sten Marcus Malva	133
<i>Mitigating Bias in Text Classification via Prompt-Based Text Transformation</i> Charmaine Barker and Dimitar Kazakov	143
<i>Towards CEFR-targeted Text Simplification for Question Adaptation</i> Luca Benedetto and Paula Buttery	150
<i>Evaluation of Pretrained and Instruction-Based Pretrained Models for Emotion Detection in Arabic Social Media Text</i> Md. Rafiul Biswas, Shima Ibrahim, Mabrouka Bessghaier and Wajdi Zaghouani	158
<i>Can LLMs Disambiguate Grounded Language? The Case of PP Attachment</i> John Blackmore and Matthew Stone	166
<i>MLDataForge: Accelerating Large-Scale Dataset Preprocessing and Access for Multimodal Foundation Model Training</i> Andrea Blasi Núñez, Lukas Paul Achatius Galke and Peter Schneider-Kamp	175
<i>The Impact of Named Entity Recognition on Transformer-Based Multi-Label Dietary Recipe Classification</i> Kemalcan Bora and Horacio Saggion	184
<i>Balancing the Scales: Addressing Gender Bias in Social Media Toxicity Detection</i> Beatriz Botella-Gil, Juan Pablo Consuegra-Ayala, Alba Bonet-Jover and Paloma Moreda-Pozo	194
<i>"Simple-Tool": A Tool for the Automatic Transformation of Spanish Texts into Easy-to-Read</i> Beatriz Botella-Gil, Isabel Espinosa-Zaragoza, Paloma Moreda Pozo and Manuel Palomar . . .	204
<i>QuARK: LLM-Based Domain-Specific Question Answering Using Retrieval Augmented Generation and Knowledge Graphs</i> Edward Burgin, Sourav Dutta and Mingxue Wang	210
<i>Classifying Emotions in Tweets from the Financial Market: A BERT-based Approach</i> Wesley Pompeu Carvalho and Norton Trevisan Roman	218
<i>Detecting Changes in Mental Health Status via Reddit Posts in Response to Global Negative Events</i> Zenan Chen, Judita Preiss and Peter A. Bath	227
<i>APIO: Automatic Prompt Induction and Optimization for Grammatical Error Correction and Text Simplification</i> Artem Chernodub, Aman Saini, Yejin Huh, Vivek Kulkarni and Vipul Raheja	234
<i>Integrating Archaic and Regional Lexicons to Improve the Readability of Old Romanian Texts</i> Madalina Chitez, Roxana Rogobete, Cristina Aura Udrea, Karla Csürös, Ana-Maria Bucur and Mihai Dascalu	240
<i>ExPe: Exact Positional Encodings for Generative Transformer Models with Extrapolating Capabilities</i> Aleksis Ioannis Datseris, Sylvia Vassileva, Ivan K. Koychev and Svetla Boytcheva	247
<i>End-to-End Deep Learning for Named Entity Recognition and Relation Extraction in Gut-Brain Axis PubMed Abstracts</i> Aleksis Ioannis Datseris, Mario Kuzmanov, Ivelina Nikolova-Koleva, Dimitar Taskov and Svetla Boytcheva	254

<i>Enabling On-Premises Large Language Models for Space Traffic Management</i> Enrique De Alba	260
<i>Top Ten from Lakhs: A Transformer-based Retrieval System for Identifying Previously Fact-Checked Claims across Multiple Languages</i> Srijani Debnath, Pritam Pal and Dipankar Das	268
<i>Evaluating Bilingual Lexicon Induction without Lexical Data</i> Michaela Denisová and Pavel Rychly	275
<i>Utilizing Large Language Models for Focused Conversational Assistants</i> Shruti Dhavalikar and Karthika Vijayan	283
<i>AntiSemRO: Studying the Romanian Expression of Antisemitism</i> Anca Dinu, Andreea C. Moldovan and Adina Marinescu	291
<i>Towards a Map of Related Words in Romance Languages</i> Liviu P. Dinu, Ana Sabina Uban, Ioan-Bogdan Iordache, Claudia Vlad, Simona Georgescu, Laurentiu Zoicas and Anca Dinu	299
<i>Decoding Emotion in Ancient Poetry: Leveraging Generative Models for Classical Chinese Sentiment Analysis</i> Quanqi Du, Loic De Langhe, Els Lefever and Veronique Hoste	306
<i>GRILE: A Benchmark for Grammar Reasoning and Explanation in Romanian LLMs</i> Marius Dumitran, Angela Dumitran and Alexandra Mihaela Danila	316
<i>PerSpaCor: Correcting Space and ZWNJ Errors in Persian Text with Transformer Models</i> Matin Ebrahimkhani and Ebrahim Ansari	325
<i>Reddit-V: A Virality Prediction Dataset and Zero-Shot Evaluation with Large Language Models</i> Samir El-amrany, Matthias R. Brust, Salima Lamsiyah and Pascal Bouvry	334
<i>Simplifications Are Absolutists: How Simplified Language Reduces Word Sense Awareness in LLM-Generated Definitions</i> Lukas Ellinger, Miriam Anschütz and Georg Groh	342
<i>Multi-LLM Text Summarization</i> Jiangnan Fang, Cheng-Tse Liu, Jieun Kim, Yash Bhedaru, Ethan Liu, Nikhil Singh, Nedim Lipka, Puneet Mathur, Nesreen K. Ahmed, Franck Dernoncourt, Ryan Rossi and Hanieh Deilamsalehy	352
<i>EDAudio: Easy Data Augmentation for Dialectal Audio</i> Lea Fischbach, Akbar Karimi, Alfred Lameli and Lucie Flek	363
<i>Authorship Verification Using Cloze Test with Large Language Models</i> Tomáš Foltýnek, Tomáš Kancko and Pavel Rychly	369
<i>A Culturally-Rich Romanian NLP Dataset from "Who Wants to Be a Millionaire?" Videos</i> Alexandru Ganea, Antonia-Adelina Popovici and Marius Dumitran	378
<i>Graph-based RAG for Low-Resource Aromanian–Romanian Translation</i> Laurentiu G. Ghetoiu and Sergiu Nisioi	388
<i>Differential Robustness in Transformer Language Models: Empirical Evaluation under Adversarial Text Attacks</i> Taniya Gidatkar, Oluwaseun Ajao and Matthew Shardlow	395

<i>An Annotation Scheme for Factuality and Its Application to Parliamentary Proceedings</i> Gili Goldin, Shira Wigderson, Ella Rabinovich and Shuly Wintner	403
<i>Can We Predict Innovation? Narrow Experts versus Competent Generalists</i> Amir Hazem and Motohashi Kazuyuki	413
<i>Arabic to Romanian Machine Translation: A Case Study on Distant Language Pairs</i> Ioan Alexandru Hirica, Stefana Arina Tabusca and Sergiu Nisioi	423
<i>BiGCAT: A Graph-Based Representation Learning Model with LLM Embeddings for Named Entity Recognition</i> Md. Akram Hossain, Abdul Aziz, Muhammad Anwarul Azim, Abu Nowshed Chy, Md Zia Ullah and Mohammad Khairul Islam	433
<i>Measuring How (Not Just Whether) VLMs Build Common Ground</i> Saki Imai, Mert Inan, Anthony B. Sicilia and Malihe Alikhani	441
<i>SiLVERScore: Semantically-Aware Embeddings for Sign Language Generation Evaluation</i> Saki Imai, Mert Inan, Anthony B. Sicilia and Malihe Alikhani	452
<i>Alignment of Historical Manuscript Transcriptions and Translations</i> Maarten Janssen, Pirooska Lendvai and Anna Jouravel	462
<i>Zero-shot OCR Accuracy of Low-Resourced Languages: A Comparative Analysis on Sinhala and Tamil</i> Nevidu Jayatilleke and Nisansa de Silva	471
<i>Detecting Gender Stereotypical Language Using Model-agnostic and Model-specific Explanations</i> Manuela Nayantara Jeyaraj and Sarah Jane Delany	481
<i>Reversing Causal Assumptions: Explainability in Online Sports Dialogues</i> Asteria Kaeberlein and Malihe Alikhani	491
<i>How LLMs Influence Perceived Bias in Journalism</i> Asteria Kaeberlein and Malihe Alikhani	501
<i>Prompting Techniques for Reducing Social Bias in LLMs through System 1 and System 2 Cognitive Processes</i> Mahammed Kamruzzaman and Gene Louis Kim	511
<i>Performance Gaps in Acted and Naturalistic Speech: Insights from Speech Emotion Recognition Strategies on Customer Service Calls</i> Lily Kawaoto, Hita Gupta, Ning Yu and Daniel Dakota	521
<i>Synthetic vs. Gold: The Role of LLM Generated Labels and Data in Cyberbullying Detection</i> Arefeh Kazemi, Sri Balaaji Natarajan Kalaivendan, Joachim Wagner, Hamza Qadeer, Kanishk Verma and Brian Davis	531
<i>FreeTxt: Analyse and Visualise Multilingual Qualitative Survey Data for Cultural Heritage Sites</i> Nouran Khallaf, Ignatius Ezeani, Dawn Knight, Paul Rayson, Mo El-Haj, John Vidler, James Davies and Fernando Alva-Manchego	541
<i>GPT-Based Lexical Simplification for Multi-Word Expressions Using Prompt Engineering</i> Sardar Khan Khayamkhani and Matthew Shardlow	546
<i>Instruction-Tuning LLaMA for Synthetic Medical Note Generation in Swedish and English</i> Lotta Kiefer, Jesujoba Alabi, Thomas Vakili, Hercules Dalianis and Dietrich Klakow	557

<i>Output Trend Analysis in Semantic Classification of Katakana Words Using a Large Language Model</i> Kazuki Kodaki and Minoru Sasaki	567
<i>Domain Knowledge Distillation for Multilingual Sentence Encoders in Cross-lingual Sentence Similarity Estimation</i> Risa Kondo, Hiroki Yamauchi, Tomoyuki Kajiware, Marie Katsurai and Takashi Ninomiya ...	572
<i>Am I Blue or Is My Hobby Counting the Teardrops? Expression Leakage in Large Language Models as a Symptom of Irrelevancy Disruption</i> Berkay Kopru, Mehrzad Mashal, Yigit Gurses, Akos Kadar, Maximilian Schmitt, Ditty Mathew, Felix Burkhardt, Florian Eyben and Björn W. Schuller	578
<i>Fusion of Object-Centric and Linguistic Features for Domain-Adapted Multimodal Learning</i> Jordan Konstantinov Kralev	587
<i>Multi-Agent Reinforcement Learning for Interactive Code Debugging with Human Feedback and Memory</i> Anjana Krishnamoorthy, Kartik Ivatury and Benyamin Ahmadnia	595
<i>Integrating Large Language Models for Comprehensive Study and Sentiment Analysis of Student Feedback</i> Jana Kuzmanova, Katerina Zdravkova and Ivan Chorbev	604
<i>Task-Oriented Dialogue Systems through Function Calling</i> Tiziano Labruna, Giovanni Bonetta and Bernardo Magnini	614
<i>When to Retrieve: Teaching LLMs to Utilize Information Retrieval Effectively</i> Tiziano Labruna, Jon Ander Campos and Gorka Azkune	623
<i>Trust but Verify: A Comprehensive Survey of Faithfulness Evaluation Methods in Abstractive Text Summarization</i> Salima Lamsiyah, Aria Nourbakhsh and Christoph Schommer	633
<i>Evaluating Large Language Models on Multiword Expressions in Multilingual and Code-Switched Contexts</i> Frances Adriana Laureano De Leon, Asim Abbas, Harish Tayyar Madabushi and Mark Lee ...	644
<i>Instruction Finetuning to Attribute Language Stage, Dialect, and Provenance Region to Historical Church Slavic Texts</i> Piroska Lendvai, Uwe Reichel, Anna Jouravel, Achim Rabus and Elena Renje	654
<i>MariATE: Automatic Term Extraction Using Large Language Models in the Maritime Domain</i> Shijie Liu, Els Lefever and Veronique Hoste	663
<i>Exploring the Usage of Knowledge Graphs in Identifying Human and LLM-Generated Fake Reviews</i> Ming Liu and Massimo Poesio	674
<i>The Evaluation of Medical Terms Complexity Using Lexical Features and Large Language Models</i> Liliya Makhmutova, Giancarlo Dondoni Salton, Fernando Perez-Tellez and Robert J. Ross	682
<i>Where and How as Key Factors for Knowledge-Enhanced Constrained Commonsense Generation</i> Ivan Martinez-Murillo, Paloma Moreda Pozo and Elena Lloret	694

<i>Forecasting Online Negativity Spikes with Multilingual Transformers for Strategic Decision-Making</i> Rowan Martnishn, Vishal Green, Varun Kadari, Shravan Athikinasetti, Zach miller, Julia Brady, Viraj Chawda and Nikhil Badlani	704
<i>C-SHAP: Collocation-Aware Explanations for Financial NLP</i> Martina Menzio, Elisabetta Fersini and Davide Paris	711
<i>Investigating Polarization in YouTube Comments via Aspect-Based Sentiment Analysis</i> Daniel Miehl, Daniel Dakota and Sandra Kübler	718
<i>From the Tractatus Logico-Philosophicus to Later Wittgenstein: An NLP-Based Comparative Analysis</i> Andreiana Mihail, Silviu-Florin Gheorghe, Andrei Fotea and Liviu P. Dinu	729
<i>Towards Intention-aligned Reviews Summarization: Enhancing LLM Outputs with Pragmatic Cues</i> Maria Miro Maestre, Robiert Sepulveda-Torres, Ernesto Luis Estevanell-Valladares, Armando Suarez Cueto and Elena Lloret	737
<i>Subtle Shifts, Significant Threats: Leveraging XAI Methods and LLMs to Undermine Language Models Robustness</i> Adrián Moreno Muñoz, L. Alfonso Ureñ-López and Eugenio Martínez Cámara	748
<i>Fast Thinking with Structured Prompts: Enabling LLM Reasoning without Chain-of-Thought Generation</i> Kirill Morozov, Liubov Chubarova and Irina Piontkovskaya	758
<i>T2Know: Analysis and Trend Platform Using the Knowledge Extracted from Scientific Texts</i> Rafael Muñoz Guillena, Manuel Palomar, Yoan Gutiérrez and Mar Bonora	767
<i>Investigating Large Language Models' (LLMs) Capabilities for Sexism Detection on a Low-Resource Language</i> Lutfiye Seda Mut Altin and Horacio Saggion	771
<i>PolyHope-M at RANLP2025 Subtask-1 Binary Hope Speech Detection: Spanish Language Classification Approach with Comprehensive Learning Using Transformer, and Traditional ML, and DL</i> Md. Julkar Naeen, Sourav Kumar Das, Sharun Akter Khushbu, Shahriar Sultan Ramit and Alaya Parven Alo	780
<i>F-LoRA-QA: Finetuning LLaMA Models with Low-Rank Adaptation for French Botanical Question Gen- eration and Answering</i> Ayoub Nainia, Régine Vignes-Lebbe, Hajar Mousannif and Jihad Zahir	787
<i>Reverse Prompting: A Novel Computational Paradigm in Schizophrenia Based on Large Language Mod- els</i> Ivan Nenchev, Christiane Montag and Sandra Anna Just	797
<i>A Survey on Small Language Models</i> Chien Van Nguyen, Xuan Shen, Ryan Aponte, Yu Xia, Samyadeep Basu, Zhengmian Hu, Jian Chen, Mihir Parmar, Sasidhar Kunapuli, Joe Barrow3, Junda Wu, Ashish Singh, Yu Wang, Jiuxiang Gu, Nesreen K. Ahmed, Nedim Lipka, Ruiyi Zhang, Xiang Chen, Tong Yu, Sungchul Kim, Hanieh Deilamsalehy, Namyong Park, Michael Rimer, Zhehao Zhang, Huanrui Yang, Puneet Mathur, Gang Wu, Franck Dernoncourt, Ryan Rossi and Thien Huu Nguyen	807
<i>Quantifying the Overlap: Attribution Maps and Linguistic Heuristics in Encoder-Decoder Machine Translation Models</i> Aria Nourbakhsh, Salima Lamsiyah and Christoph Schommer	822

<i>The Illusion of a Perfect Metric: Why Evaluating AI ‘S Words Is Harder than It Looks</i>	
Maria Paz Oliva, Adriana D. Correia, Ivan Vankov and Viktor Botev	832
<i>Multi-LLM Debiasing Framework</i>	
Deonna M. Owens, Ryan Rossi, Sungchul Kim, Tong Yu, Franck Dernoncourt, Xiang Chen, Ruiyi Zhang, Jiuxiang Gu, Hanieh Deilamsalehy and Nedim Lipka	843
<i>Toward Quantum-Enhanced Natural Language Understanding: Sarcasm and Claim Detection with QLSTM</i>	
Pritam Pal and Dipankar Das	854
<i>Legal Terminology Extraction in Spanish: Gold-standard Generation and LLM Evaluation</i>	
Lucia Palacios Palacios, Beatriz Guerrero García, Patricia Martín Chozas and Elena Montiel Ponsoda	860
<i>Benchmarking Item Difficulty Classification in German Vocational Education and Training</i>	
Alonso Palomino and Benjamin Paassen	870
<i>Isolating LLM Performance Gains in Pre-training versus Instruction-tuning for Mid-resource Languages: The Ukrainian Benchmark Study</i>	
Yurii Paniv	876
<i>Evaluating LLMs on Deceptive Text across Cultures</i>	
Katerina Papantoniou, Panagiotis Papadakos and Dimitris Plexousakis	884
<i>Annotating Hate Speech towards Identity Groups</i>	
Donnie Parent, Nina Georgiades, Charvi Mishra, Khaled Mohammed and Sandra Kübler	894
<i>On the Interaction of Identity Hate Classification and Data Bias</i>	
Donnie Parent, Nina Georgiades, Charvi Mishra, Khaled Mohammed and Sandra Kübler	900
<i>Financial News as a Proxy of European Central Bank Interest Rate Adjustments</i>	
Davide Paris, Martina Menzio and Elisabetta Fersini	907
<i>Generating and Analyzing Disfluency in a Code-Mixed Setting</i>	
Aryan Paul, Tapabrata Mondal, Dipankar Das and Sivaji Bandyopadhyay	915
<i>A Low-Resource Speech-Driven NLP Pipeline for Sinhala Dyslexia Assistance</i>	
Peshala Sandali Perera and Deshan Koshala Sumanathilaka	925
<i>Evaluating Transliteration Ambiguity in Adhoc Romanized Sinhala: A Dataset for Transliteration Disambiguation</i>	
Sandun Sameera Perera and Deshan Koshala Sumanathilaka	934
<i>Detecting Deception in Disinformation across Languages: The Role of Linguistic Markers</i>	
Alba Perez-Montero, Silvia Gargova, Elena Lloret and Paloma Moreda Pozo	943
<i>Enhancing Transformer-Based Rerankers with Synthetic Data and LLM-Based Supervision</i>	
Dimitar Peshevski, Kiril Blazhevski, Martin Popovski and Gjorgji Madjarov	953
<i>Q&A-LF : A French Question-Answering Benchmark for Measuring Fine-Grained Lexical Knowledge</i>	
Alexander Petrov, Alessandra Thais Mancas, Viviane Binet, Antoine Venant, Francois Lareau, Yves Lepage and Phillippe Langlais	962

<i>Analysis of Vocabulary and Subword Tokenization Settings for Optimal Fine-tuning of MT: A Case Study of In-domain Translation</i>	
Javad Pourmostafa Roshan Sharami, Dimitar Shterionov and Pieter Spronck	970
<i>LLM-based Embedders for Prior Case Retrieval</i>	
Damith Premasiri, Tharindu Ranasinghe and Ruslan Mitkov	980
<i>Exploiting Primacy Effect to Improve Large Language Models</i>	
Bianca Raimondi and Maurizio Gabbrielli	989
<i>Alankaar: A Dataset for Figurativeness Understanding in Bangla</i>	
Geetanjali Rakshit and Jeffrey Flanigan	998
<i>ASQ: Automatically Generating Question-Answer Pairs Using AMRs</i>	
Geetanjali Rakshit and Jeffrey Flanigan	1003
<i>Multi-LLM Verification for Question Answering under Conflicting Contexts</i>	
Geetanjali Rakshit and Jeffrey Flanigan	1012
<i>Comparative Analysis of Human and Large Language Model Performance in Pharmacology Multiple-Choice Questions</i>	
Ricardo Rodriguez, StÅ©phane Huet, Benoit Favre and Mickael Rouvier	1022
<i>Enhancing Textual Understanding: Automated Claim Span Identification in English, Hindi, Bengali, and CodeMix</i>	
Rudra Roy, Pritam Pal, Dipankar Das, Saptarshi Ghosh and Biswajit Paul	1030
<i>Detecting Fake News in the Era of Language Models</i>	
Muhammad Irfan Fikri Sabri, Hansi Hettiarachchi and Tharindu Ranasinghe	1036
<i>Cyberbullying Detection via Aggression-Enhanced Prompting</i>	
Aisha Saeid, Anu Sabu, Girish Koushik, Ferrante Neri and Diptesh Kanojia	1044
<i>Lingdex.org: Leveraging LLMs to Structure and Explore Linguistic Olympiad Puzzles for Learning and Teaching Linguistics</i>	
Jonathan Sakunkoo and Annabella Sakunkoo	1053
<i>When Does Language Transfer Help? Sequential Fine-Tuning for Cross-Lingual Euphemism Detection</i>	
Julia Sammartino, Libby Barak, Jing Peng and Anna Feldman	1058
<i>Modelling the Relative Contributions of Stylistic Features in Forensic Authorship Attribution</i>	
G. Çağatay Sat, John Blake and Evgeny Pyshkin	1066
<i>The Hidden Cost of Structure: How Constrained Decoding Affects Language Model Performance</i>	
Maximilian Schall and Gerard de Melo	1074
<i>A Question-Answering Based Framework/Metric for Evaluation of Newspaper Article Summarization</i>	
Vasanth Seemakurthy, Shashank Sundar, Siddharth Arvind, Siddhant Jagdish and Ashwini M. Joshi	1085
<i>Efficient Financial Fraud Detection on Mobile Devices Using Lightweight Large Language Models</i>	
Lakpriya Senevirathna and Deshan Koshala Sumanathilaka	1090
<i>Contextual Cues in Machine Translation: Investigating the Potential of Multi-Source Input Strategies in LLMs and NMT Systems</i>	
Lia Shahnazaryan, Patrick Simianer and Joern Wuebker	1099

<i>Exposing Pink Slime Journalism: Linguistic Signatures and Robust Detection against LLM-Generated Threats</i>	
Sadat Shahriar, Navid Ayoobi, Arjun Mukherjee, Mostafa Musharrat and Sai Vishnu Vamsi Sena-gasetty	1109
<i>The Erosion of LLM Signatures: Can We Still Distinguish Human and LLM-Generated Scientific Ideas after Iterative Paraphrasing?</i>	
Sadat Shahriar, Navid Ayoobi and Arjun Mukherjee	1118
<i>Deep Language Geometry: Constructing a Metric Space from LLM Weights</i>	
Maksym Shamrai and Vladyslav Hamolia	1127
<i>Cross-Lingual Fact Verification: Analyzing LLM Performance Patterns across Languages</i>	
Hanna Shcharbakova, Tatiana Anikina, Natalia Skachkova and Josef van Genabith	1137
<i>ESAQueryRank: Ranking Query Interpretations for Document Retrieval Using Explicit Semantic Analysis</i>	
Aviyeet Shil and Wei Jin	1148
<i>Personalized Author Obfuscation with Large Language Models</i>	
Mohammad Shokri, Sarah Ita Levitan and Rivka Levitan	1153
<i>Bulgarian Event Extraction with LLMs</i>	
Kiril Simov, Nikolay Paev, Petya Osenova and Stefan Marinov	1163
<i>FigCaps-HF: A Figure-to-Caption Generative Framework and Benchmark with Human Feedback</i>	
Ashish Singh, Ashutosh Singh, Prateek Agarwal, Zixuan Huang, Arpita Singh, Tong Yu, Sungchul Kim, Victor Soares Bursztyn, Nesreen K. Ahmed, Puneet Mathur, Erik Learned-Miller, Franck Dernon-court and Ryan Rossi	1172
<i>LLM Compression: How Far Can We Go in Balancing Size and Performance?</i>	
Sahil Sk, Debashish Dhal, Sonal Khosla, Akash Dhaka, Shantipriya Parida, Sk Shahid, Sambit Shekhar, Dilip Prasad and Ondrej Bojar	1183
<i>Pushing the (Generative) Envelope: Measuring the Effect of Prompt Technique and Temperature on the Generation of Model-based Systems Engineering Artifacts</i>	
Erin Smith Crabb, Cedric Bernard, Matthew Jones and Daniel Dakota	1188
<i>Dutch CrowS-Pairs: Adapting a Challenge Dataset for Measuring Social Biases in Language Models for Dutch</i>	
Elza Strazda and Gerasimos Spanakis	1195
<i>The Challenge of Performing Ontology-driven Entity Extraction in Real-world Unstructured Textual Data from the Domain of Dementia</i>	
Sumaiya Suravee, Carsten Oliver Schmidt and Kristina Yordanova	1205
<i>Recognizing the Structure and Content of Hungarian Civil Registers</i>	
Kata Ágnes Szűcs, Noémi Vadász and Zsolt Béla Záros	1215
<i>Optimism, Pessimism, and the Language between: Model Interpretability and Psycholinguistic Profiling</i>	
Stefana Arina Tabusca and Liviu P. Dinu	1224
<i>Demographic Features for Annotation-Aware Classification</i>	
Narjes Tahaei and Sabine Bergler	1232

<i>Exploring the Performance of Large Language Models for Event Detection and Extraction in the Health Domain</i>	
Hristo Tanev, Nicolas Stefanovitch, Tomáš Harmatha and Diana F. Sousa	1237
<i>Leveraging LLaMa for Abstractive Text Summarisation in Malayalam: An Experimental Study</i>	
Hristo Tanev, Anitha S. Pillai and Revathy V. R.	1248
<i>Building a Clean Bartangi Language Corpus and Training Word Embeddings for Low-Resource Language Modeling</i>	
Warda Tariq, Victor Popov and Vasilii Gromov	1256
<i>A Deep Dive into Multi-Head Attention and Multi-Aspect Embedding</i>	
Maryam Teimouri, Jenna Kanerva and Filip Ginter.	1263
<i>A Linguistically-informed Comparison between Multilingual BERT and Language-specific BERT Models: The Case of Differential Object Marking in Romanian</i>	
Maria Tepei and Jelke Bloem	1271
<i>PoliStance-TR: A Dataset for Turkish Stance Detection in Political Domain</i>	
Muhammed Cihat Unal, Yasemin Sarkin, Alper Karamanlioglu and Berkan Demirel	1282
<i>Towards Safer Hebrew Communication: A Dataset for Offensive Language Detoxification</i>	
Natalia Vanetik, Lior Liberov, Marina Litvak and Chaya Liebeskind	1289
<i>AIDEN: Automatic Speaker Notes Creation and Navigation for Enhancing Online Learning Experience</i>	
Stalin Varanasi, Umer Butt, Guenter Neumann and Josef van Genabith.	1299
<i>Using LLMs for Multilingual Clinical Entity Linking to ICD-10</i>	
Sylvia Vassileva, Ivan K. Koychev and Svetla Boytcheva	1304
<i>Aspect–Sentiment Quad Prediction with Distilled Large Language Models</i>	
Filippos Karolos Ventirozos, Peter Appleby and Matthew Shardlow	1309
<i>SENTimental - a Simple Multilingual Sentiment Annotation Tool</i>	
John Vidler, Paul Rayson and Dawn Knight	1320
<i>Anonymise: A Tool for Multilingual Document Pseudonymisation</i>	
Rinalds Vīksna and Inguna Skadina	1327
<i>Revealing Gender Bias in Language Models through Fashion Image Captioning</i>	
Maria Villalba-Oses, Victoria Muñoz-Garcia and Juan Pablo Consuegra-Ayala	1333
<i>Benchmarking Korean Idiom Understanding: A Comparative Analysis of Local and Global Models</i>	
Xiaonan Wang, Seoyoon Park and Hansaem Kim	1341
<i>TinyMentalLLMs Enable Depression Detection in Chinese Social Media Texts</i>	
JINYUAN XU, Tian LAN, Mathieu Valette, Pierre Magistry and LEI LI	1352
<i>Prompt Engineering for Nepali NER: Leveraging Hindi-Capable LLMs for Low-Resource Languages</i>	
Dipendra Yadav, Sumaiya Suravee, Stefan Kemnitz, Tobias Strauss and Kristina Yordanova	1364
<i>Seeing, Signing, and Saying: A Vision-Language Model-Assisted Pipeline for Sign Language Data Acquisition and Curation from Social Media</i>	
Shakib Yazdani, Yasser HAMIDULLAH, Cristina España-Bonet and Josef van Genabith	1374

<i>Visual Priming Effect on Large-scale Vision Language Models</i>	
Daiki Yoshida, Haruki Sakajo, Kazuki Hayashi, Yusuke Sakai, Hidetaka Kamigaito, Katsuhiko Hayashi and Taro Watanabe	1385
<i>From Courtroom to Corpora: Building a Name Entity Corpus for Urdu Legal Texts</i>	
Adeel Zafar, Sohail Ashraf and Slawomir Nowaczyk	1396
<i>EmoHopeSpeech: An Annotated Dataset of Emotions and Hope Speech in English and Arabic</i>	
Wajdi Zaghouani and Md. Rafiul Biswas	1406
<i>An Annotated Corpus of Arabic Tweets for Hate Speech Analysis</i>	
Wajdi Zaghouani and Md. Rafiul Biswas	1413
<i>Strategies for Efficient Retrieval-augmented Generation in Clinical Domains with RAPTOR: A Benchmarking Study</i>	
Xumou Zhang, Qixuan Hu, Jinman Kim and Adam G. Dunn	1420
<i>LLM-Based Product Recommendation with Prospect Theoretic Self Alignment Strategy</i>	
Manying Zhang, Zehua Cheng and Damien Nouvel	1430
<i>Branching Out: Exploration of Chinese Dependency Parsing with Fine-tuned Large Language Models</i>	
He Zhou, Emmanuele Chersoni and Yu-Yin Hsu	1437