



**Bulgarian Academy of Sciences  
Institute of Information and  
Communication Technologies**

## PROGRAMME

of the 15<sup>th</sup> International Conference on

**Recent Advances in Natural Language Processing**

**RANLP-2025**



**8-10 September, 2025**

**CHERNO MORE Hotel, Varna, Bulgaria**



КП-06-МНФ/11, 09.06.2025



IRIS.AI



# RANLP 2025 Programme

## Day 1 (8 September 2025, Monday)

9.00-9.15 **Opening**

9.15-10.15 **Keynote speech 1** (Room: Chernomore hall) **Anna Rogers**

*Large language models and factuality*

(50 min presentation; 10 min Q&A)

Introduction Ruslan Mitkov

10.15-11.05 **Parallel sessions 1-4 (regular papers: 20 min presentation; 5 min Q&A)**

### **Parallel session 01: Cross-Lingual Transfer and Prompting Strategies**

Room - Chernomore hall; Chair - Saad Ezzini

- 10.15-10.40 When Does Language Transfer Help? Sequential Fine-Tuning for Cross-Lingual Euphemism Detection (Julia Sammartino, Libby Barak, Jing Peng and Anna Feldman)
- 10.40-11.05 Prompt Engineering for Nepali NER: Leveraging Hindi-Capable LLMs for Low-Resource Languages (Dipendra Yadav, Sumaiya Suravee, Stefan Kemnitz, Tobias Strauss and Kristina Yordanova)

### **Parallel session 02: Machine Translation**

Room - Varna hall; Chair - Alicia Picazo

- 10.15-10.40 Advancing Clinical Translation in Nepali with Fine-Tuned Multilingual Models (Benyamin Ahmadnia, Sumaiya Shaikh, Bibek Poudel, Shazan Mohammed and Sahar Hooshmand)
- 10.40-11.05 Quantifying the Overlap: Attribution Maps and Linguistic Heuristics in Encoder-Decoder Machine Translation Models (Aria Nourbakhsh, Salima Lamsiyah and Christoph Schommer)

### **Parallel session 03: Aspect-Based Sentiment Analysis**

Room - Odessos Hall; Chair - Eugenio Martínez Camara

- 10.15-10.40 Investigating Polarization in YouTube Comments via Aspect-Based Sentiment Analysis (Daniel Miebling, Daniel Dakota and Sandra Kübler)
- 10.40-11.05 Aspect-Sentiment Quad Prediction with Distilled Large Language Models (Filippos Karolos Ventirozos, Peter Appleby and Matthew Shardlow)

Check related online presentations or papers in Proceedings: (regular) TinyMentalLLMs Enable Depression Detection in Chinese Social Media Texts (Jinyuan Xu, Tian Lan, Mathieu Valette, Pierre Magistry and Lei Li)

#### **Parallel session 04: Sentiment analysis and generation**

Room - Galata hall; Chair - Max-Müller-Eberstein

- 10.15-10.40 Where and How as Key Factors for Knowledge-Enhanced Constrained Commonsense Generation (Ivan Martinez-Murillo, Paloma Moreda Pozo and Elena Lloret)
- 10.40-11.05 Decoding Emotion in Ancient Poetry: Leveraging Generative Models for Classical Chinese Sentiment Analysis (Quanqi Du, Loic De Langhe, Els Lefever and Veronique Hoste)

11.05-11.35 **Coffee break**

11.35-12.50 **Parallel sessions 5-8 (regular papers: 20 min presentation; 5 min Q&A)**

#### **Parallel session 05: NER, Terminology, Idiom Extraction**

Room - Chernomore hall; Chair - Paul Rayson

- 11:35-12:00 The Impact of Named Entity Recognition on Transformer-Based Multi-Label Dietary Recipe Classification (Kemalcan Bora and Horacio Saggion)
- 12:00-12:25 MariATE: Automatic Term Extraction Using Large Language Models in the Maritime Domain (Shijie Liu, Els Lefever and Veronique Hoste)
- 12:25-12:50 Benchmarking Korean Idiom Understanding: A Comparative Analysis of Local and Global Models (Xiaonan Wang, Seoyoon Park and Hansaem Kim)

Check related recorded presentations or papers in Proceedings:

- (regular) Evaluating Transliteration Ambiguity in Adhoc Romanized Sinhala: A Dataset for Transliteration Disambiguation (Sandun Sameera Perera and Deshan Koshala Sumanathilaka)
- (regular) BiGCAT: A Graph-Based Representation Learning Model with LLM Embeddings for Named Entity Recognition (Md. Akram Hossain, Abdul Aziz, Muhammad Anwarul Azim, Abu Nowshed Chy, Md Zia Ullah and Mohammad Khairul Islam)

#### **Parallel session 06: Lexical and World Knowledge**

Room - Varna hall; Chair - Petya Osenova

- 11:35-12:00 GRILE: A Benchmark for Grammar Reasoning and Explanation in Romanian LLMs (Marius Dumitran, Angela Dumitran and Alexandra Mihaela Danila)
- 12:00-12:25 Q&A-LF: A French Question-Answering Benchmark for Measuring Fine-Grained Lexical Knowledge (Alexander Petrov, Alessandra Thais Mancas, Viviane Binet, Antoine Venant, Francois Lareau, Yves Lepage and Phillippe Langlais)
- 12:25-12:50 A Culturally-Rich Romanian NLP Dataset from "Who Wants to Be a Millionaire?" Videos (Alexandru Ganea, Antonia-Adelina Popovici and Marius Dumitran)

### **Parallel session 07: VLM -- Sign Language and Figure-to-Caption**

Room - Odessos Hall; Chair - Sandra Kübler

- 11:35-12:00 SILVERScore: Semantically-Aware Embeddings for Sign Language Generation Evaluation (Saki Imai, Mert Inan, Anthony B. Sicilia and Malihe Alikhani)
- 12:00-12:25 Seeing, Signing, and Saying: A Vision-Language Model-Assisted Pipeline for Sign Language Data Acquisition and Curation from Social Media (Shakib Yazdani, Yasser Hamidullah, Cristina España-Bonet and Josef van Genabith)
- 12:25-12:50 FigCaps-HF: A Figure-to-Caption Generative Framework and Benchmark with Human Feedback (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 Deroncourt and Ryan Rossi)

### **Parallel session 08: Multi-Language Spaces and Summarisation**

Room - Galata hall; Chair - Salima Lamsiya

- 11:35-12:00 Deep Language Geometry: Constructing a Metric Space from LLM Weights (Maksym Shamrai and Vladyslav Hamolia)
- 12:00-12:25 Multi-LLM Text Summarization (Jiangnan Fang, Cheng-Tse Liu, Jieun Kim, Yash Bhedaru, Ethan Liu, Nikhil Singh, Nedim Lipka, Puneet Mathur, Nesreen K. Ahmed, Franck Deroncourt, Ryan Rossi and Hanieh Deilamsalehy)

Check related remoted pre-recorded presentations or papers in Proceedings: (regular) Exposing Pink Slime Journalism: Linguistic Signatures and Robust Detection against LLM-Generated Threats (Sadat Shahriar, Navid Ayoobi, Arjun Mukherjee, Mostafa Musharrat and Sai Vishnu Vamsi Senagasetty)

12.50-14.50 pm **Lunch break**

14.50-15.35 **Diamond sponsor presentation** (Mind Bridge AI)

15.35-16.35 **Student Research Workshop Talks (regular papers: 20 min presentation; 5 min Q&A)**

Room - Chernomore hall; Chair - Ivelina Nikolova

- 15.35-15.40 Opening
- 15.40 - 16.05 Automated classification of causal relations. Evaluating different LLM performances. (Giacomo Magnifico)
- 16.05 - 16.30 Study on Automatic Punctuation Restoration in Bilingual Broadcast Stream (Martin Polacek)

## 15.35-16.35 **Poster session 1 and Demo session 1**

Chair: Dimitar Kazakov

### **Demos:**

1. T2Know: Analysis and Trend Platform Using the Knowledge Extracted from Scientific Texts (Rafael Muñoz Guillena, Manuel Palomar, Yoan Gutiérrez and Mar Bonora)
2. FreeTxt: Analyse and Visualise Multilingual Qualitative Survey Data for Cultural Heritage Sites (Nouran Khallaf, Ignatius Ezeani, Dawn Knight, Paul Rayson, Mo El-Haj, John Vidler, James Davies and Fernando Alva-Manchego)
3. Anonymise: A Tool for Multilingual Document Pseudonymisation (Rinalds Vīksna and Inguna Skadiņa)
4. "Simple-Tool": A Tool for the Automatic Transformation of Spanish Texts into Easy-to-Read (Beatriz Botella-Gil, Isabel Espinosa-Zaragoza, Paloma Moreda Pozo and Manuel Palomar)

### **Posters:**

1. On the Limitations of Large Language Models (LLMs): False Attribution (Tosin Adewumi, Nudrat Habib, Lama Alkhaled and Elisa Barney)
2. ESAQueryRank: Ranking Query Interpretations for Document Retrieval Using Explicit Semantic Analysis (Avijeet Shil and Wei Jin)
3. Recognizing the Structure and Content of Hungarian Civil Registers (Kata Ágnes Szűcs, Noémi Vadász and Zsolt Béla Záros)
4. Output Trend Analysis in Semantic Classification of Katakana Words Using a Large Language Model (Kazuki Kodaki and Minoru Sasaki)
5. Exploiting Primacy Effect to Improve Large Language Models (Bianca Raimondi and Maurizio Gabbrielli)
6. Comparative Analysis of Human and Large Language Model Performance in Pharmacology Multiple-Choice Questions (Ricardo Rodriguez, Stéphane Huet, Benoit Favre and Mickael Rouvier)
7. Evaluating Bilingual Lexicon Induction without Lexical Data (Michaela Denisová and Pavel Rychly)
8. A Framework for Fine-Tuning LLMs Using Heterogeneous Feedback (Ryan Aponte, Ryan A. Rossi, Shunan Guo, Franck Dernoncourt, Tong Yu, Xiang Chen, Subrata Mitra and Nedim Lipka)
9. How LLMs Influence Perceived Bias in Journalism (Asteria Kaeberlein and Malihe Alikhani)
10. Integrating Large Language Models for Comprehensive Study and Sentiment Analysis of Student Feedback (Jana Kuzmanova, Katerina Zdravkova and Ivan Chorbev)
11. Strategies for Efficient Retrieval-augmented Generation in Clinical Domains with RAPTOR: A Benchmarking Study (Xumou Zhang, Qixuan Hu, Jinman Kim and Adam G. Dunn)
12. Legal Terminology Extraction in Spanish: Gold-standard Generation and LLM Evaluation (Lucia Palacios Palacios, Beatriz Guerrero García, Patricia Martín Chozas and Elena Montiel Ponsoda)
13. PoliStance-TR: A Dataset for Turkish Stance Detection in Political Domain (Muhammed Cihat Unal, Yasemin Sarkın, Alper Karamanlioglu and Berkan Demirel)

14. Integrating Archaic and Regional Lexicons to Improve the Readability of Old Romanian Texts (Madalina Chitez, Roxana Rogobete, Cristina Aura Udrea, Karla Csürös, Ana-Maria Bucur and Mihai Dascalu)
15. Using LLMs for Multilingual Clinical Entity Linking to ICD-10 (Sylvia Vassileva, Ivan K. Koychev and Svetla Boytcheva)
16. Towards Safer Hebrew Communication: A Dataset for Offensive Language Detoxification (Natalia Vanetik, Lior Liberov, Marina Litvak and Chaya Liebeskind)
17. Classifying Emotions in Tweets from the Financial Market: A BERT-based Approach (Wesley Pompeu Carvalho and Norton Trevisan Roman)
18. Annotating Hate Speech towards Identity Groups (Donnie Parent, Nina Georgiades, Charvi Mishra, Khaled Mohammed and Sandra Kübler)
19. On the Interaction of Identity Hate Classification and Data Bias (Donnie Parent, Nina Georgiades, Charvi Mishra, Khaled Mohammed and Sandra Kübler)
20. From Posts to Predictions: A User-Aware Framework for Faithful and Transparent Detection of Mental Health Risks on Social Media (Hessam Amini and Leila Kosseim)
21. LLM-Based Product Recommendation with Prospect Theoretic Self Alignment Strategy (Manying Zhang, Zehua Cheng and Damien Nouvel)
22. Multilingual Pre-training Meets Supervised Neural Machine Translation: A Reproducible Evaluation on English–French and Finnish Translation (Benyamin Ahmadnia, Yeswanth Soma and Hossein Sarrafzadeh)
23. Enhancing Textual Understanding: Automated Claim Span Identification in English, Hindi, Bengali, and CodeMix (Rudra Roy, Pritam Pal, Dipankar Das, Saptarshi Ghosh and Biswajit Paul)
24. Toward Quantum-Enhanced Natural Language Understanding: Sarcasm and Claim Detection with QLSTM (Pritam Pal and Dipankar Das)
25. Leveraging LLaMa for Abstractive Text Summarisation in Malayalam: An Experimental Study (Hristo Tanev, Anitha S. Pillai and Revathy V. R)
26. EDAudio: Easy Data Augmentation for Dialectal Audio (Lea Fischbach, Akbar Karimi, Alfred Lameli and Lucie Flek)
27. Lingdex.org: Leveraging LLMs to Structure and Explore Linguistic Olympiad Puzzles for Learning and Teaching Linguistics (Jonathan Sakunkoo and Annabella Sakunkoo)
28. Arabic to Romanian Machine Translation: A Case Study on Distant Language Pairs (Ioan Alexandru Hirica, Stefana Arina Tabusca and Sergiu Nisioi)
29. Towards a Map of Related Words in Romance Languages (Liviu P. Dinu, Ana Sabina Uban, Ioan-Bogdan Iordache, Claudia Vlad, Simona Georgescu, Laurentiu Zoicas and Anca Dinu)
30. HoloBERT: Pre-Trained Transformer Model for Historical Narratives (Isuri Anuradha, Le An Ha and Ruslan Mitkov)
31. Forecasting Online Negativity Spikes with Multilingual Transformers for Strategic Decision-Making (Rowan Martnishn, Vishal Green, Varun Kadari, Shravan Athikinasetti, Zach Miller, Julia Brady, Viraj Chawda and Nikhil Badlani)

16.35-17.00: **Coffee break**

17.00-17.20: **Gold sponsor presentation** (Graphwise)

## **Day 2 (9 September, Tuesday)**

8:30-12:00 **Excursion**

12.00-14.00 **Lunch break**

14.00-15.00 **Keynote speech 2** (Room: Chernomore hall) **Roberto Navigli**

*Do Large Language Models Understand Word Meanings?*

(50 min presentation; 10 min Q&A)

Introduction Ruslan Mitkov

15.00-16.00 **Parallel sessions 9-12 (short papers: 15 min presentation; 5 min Q&A)**

### **Parallel session 09: Multi-LLM and LLM-as-a-Judge Scenarios**

Room - Chernomore hall; Chair - Salima Lamsiya

15.00-15.20 Multi-LLM Debiasing Framework (Deonna M. Owens, Ryan Rossi, Sungchul Kim, Tong Yu, Franck Dernoncourt, Xiang Chen, Ruiyi Zhang, Jiuxiang Gu, Hanieh Deilamsalehy and Nedim Lipka)

15.20-15.40 Multi-LLM Verification for Question Answering under Conflicting Contexts (Geetanjali Rakshit and Jeffrey Flanigan)

15.40-16.00 The Illusion of a Perfect Metric: Why Evaluating AI'S Words Is Harder than It Looks (Maria Paz Oliva, Adriana D. Correia, Ivan Vankov and Viktor Botev)

### **Parallel session 10: Annotated Datasets: Creative Language and Factuality**

Room - Varna hall; Chair - Wajdi Zaghouani

15.00-15.20 KoWit-24: A Richly Annotated Dataset of Wordplay in News Headlines (Alexander Baranov, Anna Palatkina, Yulia Makovka and Pavel Braslavski)

15.20-15.40 Alankaar: A Dataset for Figurativeness Understanding in Bangla (Geetanjali Rakshit)

15.40-16.00 An Annotation Scheme for Factuality and Its Application to Parliamentary Proceedings (Gili Goldin, Shira Wigderson, Ella Rabinovich and Shuly Wintner)

### **Parallel session 11: Natural Language Generation and Gender Bias**

Room - Odessos Hall; Chair - Beatriz Botella

- 15.00-15.20 AIDEN: Automatic Speaker Notes Creation and Navigation for Enhancing Online Learning Experience (Stalin Varanasi, Umer Butt, Guenter Neumann and Josef van Genabith)
- 15.20-15.40 Reverse Prompting: A Novel Computational Paradigm in Schizophrenia Based on Large Language Models (Ivan Nenchev, Christiane Montag and Sandra Anna Just)
- 15.40-16.00 Revealing Gender Bias in Language Models through Fashion Image Captioning (Maria Villalba-Oses, Victoria Muñoz-Garcia and Juan Pablo Consuegra-Ayala)

### **Parallel session 12: Event extraction, named entity recognition and LLMs under threat**

Room - Galata hall; Chair - Hansi Hettiarachchi

- 15.00-15.20 Event Extraction for Bulgarian with LLMs (Kiril Simov, Nikolay Paev, Petya Osenova and Stefan Marinov)
- 15.20-15.40 Beyond Methods and Datasets Entities: Introducing SH-NER for Hardware and Software Entity Recognition in Scientific Text (Aftab Anjum, Nimra Maqbool and Ralf Krestel)
- 15.40-16.00 Subtle Shifts, Significant Threats: Leveraging XAI Methods and LLMs to Undermine Language Models Robustness (Adrián Moreno Muñoz, L. Alfonso Ureña-López and Eugenio Martínez Cámara)

16.00-16.20 **Coffee break**

16.20-16:40 **Gold Sponsor presentation ([Iris.AI](#))**

16:45-17:45 **Keynote speech 3** (Room: Chernomore hall) **Eneko Agirre**

*LLMs and low-resource languages*

(50 min presentation; 10 min Q&A)

Introduction Galia Angelova

17:45-18.30 **Panel discussion ‘NLP, LLMs, and AI: Where Are We Heading’?**

With the participation of Roberto Navigi, Eneko Agirre, Preslav Nakov, Manuel Palomar, Sandra Kübler, Saad Ezzini and Tharindu Ranasinghe

Moderator: Ruslan Mitkov

### 18.30-19.00 SEPLN tribute

Brief Presentation: Rafael Muñoz Guillena (President SEPLN)

Tributes to SEPLN members and executives

### 20.30-... Gala dinner

**Dress code:** smart or smart casual. No shorts, no flip-flops please.

## Day 3 (10 September, Wednesday)

9.00-10.00 **Keynote speech 4** (Room: Chernomore hall) **Preslav Nakov**

*Towards Truly Open, Language-Specific, Safe, Factual, and Specialised Large Language Models*

(50 min presentation; 10 min Q&A)

Introduction: Galia Angelova

10.00-10:50 **Parallel session 13 (regular papers: 20 min presentation; 5 min Q&A)**

### Parallel session 13: Sentiment Analysis and Bias Detection

Room - Chernomore hall; Chair - Maram Alharbi

10.00-10.25 Am I Blue or Is My Hobby Counting the Teardrops? Expression Leakage in Large Language Models as a Symptom of Irrelevancy Disruption (Berkay Kopru, Mehrzad Mashal, Yigit Gurses, Akos Kadar, Maximilian Schmitt, Ditty Mathew, Felix Burkhardt, Florian Eyben and Björn W. Schuller)

10.25-10.50 Dutch CrowS-Pairs: Adapting a Challenge Dataset for Measuring Social Biases in Language Models for Dutch (Elza Strazda and Gerasimos Spanakis)

Check related remote pre-recorded presentations or papers in Proceedings: (short) AntiSemRO: Studying the Romanian Expression of Antisemitism (Anca Dinu, Andreea C. Moldovan and Adina Marincea)

10.00-10:40 **Parallel sessions 14-16 (short papers: 15 min presentation; 5 min Q&A)**

### Parallel session 14: Summarisation

Room - Varna hall; Chair - Mathew Shardlow

10.00-10.20 Towards Intention-aligned Reviews Summarization: Enhancing LLM Outputs with Pragmatic Cues (María Miró Maestre, Robiert Sepúlveda-Torres, Ernesto Luis Estevanell-Valladares, Armando Suárez Cueto and Elena Lloret)

10.20-10.40 Trust but Verify: A Comprehensive Survey of Faithfulness Evaluation Methods in Abstractive Text Summarization (Salima Lamsiyah, Aria Nourbakhsh and Christoph Schommer)

### **Parallel session 15: Advances in Machine Translation**

Room - Odessos Hall; Chair - Marie Escribe

- 10.00-10.20 Analysis of Vocabulary and Subword Tokenization Settings for Optimal Fine-tuning of MT: A Case Study of In-domain Translation (Javad Pourmostafa Roshan Sharami, Dimitar Shterionov and Pieter Spronck)
- 10.20-10.40 Contextual Cues in Machine Translation: Investigating the Potential of Multi-Source Input Strategies in LLMs and NMT Systems (Lia Shahnazaryan, Patrick Simianer and Joern Wuebker)

### **Parallel session 16: Constrained Output Generation**

Room - Galata hall; Chair - Ivelina Nikolova

- 10.00-10.20 The Hidden Cost of Structure: How Constrained Decoding Affects Language Model Performance (Maximilian Schall and Gerard de Melo)
- 10.20-10.40 Branching Out: Exploration of Chinese Dependency Parsing with Fine-tuned Large Language Models (He Zhou, Emmanuele Chersoni and Yu-Yin Hsu)

Check related online presentations or papers in Proceedings: (short) PerSpaCor: Correcting Space and ZWNJ Errors in Persian Text with Transformer Models (Matin Ebrahimkhani and Ebrahim Ansari)

10.40-11:20 **Coffee break**

11:20-12.10 **Parallel session 17 (regular papers: 20 min presentation; 5 min Q&A)**

### **Parallel session 17: Small Language Models and Model Compression Effects**

Room - Chernomore hall; Chair - Isuri Anuradha

- 11:20-11.45 A Survey on Small Language Models (Chien Van Nguyen, Xuan Shen, Ryan Aponte, Yu Xia, Samyadeep Basu, Zhengmian Hu, Jian Chen, Mihir Parmar, Sasidhar Kunapuli, Joe Barrow<sup>3</sup>, 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 Deroncourt, Ryan Rossi and Thien Huu Nguyen)
- 11.45-12.10 Enabling On-Premises Large Language Models for Space Traffic Management (Enrique De Alba)

Check related pre-recorded presentations or papers in Proceedings: (short) Efficient Financial Fraud Detection on Mobile Devices Using Lightweight Large Language Models (Lakpriya Senevirathna and Deshan Koshala Sumanathilaka)

11:20-12.20 **Parallel sessions 18-20 (short papers: 15 min presentation; 5 min Q&A)**

### **Parallel session 18: Synthetic Data Generation**

Room - Varna hall; Chair - Alba Bonet

- 11.20-11.40 Instruction-Tuning LLaMA for Synthetic Medical Note Generation in Swedish and English (Lotta Kiefer, Jesujoba Alabi, Thomas Vakili, Hercules Dalianis and Dietrich Klakow)
- 11.40-12.00 Generating and Analyzing Disfluency in a Code-Mixed Setting (Aryan Paul, Tapabrata Mondal, Dipankar Das and Sivaji Bandyopadhyay)
- 12.00-12.20 Investigating Large Language Models' (LLMs) Capabilities for Sexism Detection on a Low-Resource Language (Lutfiye Seda Mut Altin and Horacio Saggion)

### **Parallel session 19: IR -- Document and Sentence Alignment**

Room - Odessos Hall; Chair - Dimitar Kazakov

- 11.20-11.40 Alignment of Historical Manuscript Transcriptions and Translations (Maarten Janssen, Piroska Lendvai and Anna Jouravel)
- 11.40-12.00 Top Ten from Lakhs: A Transformer-based Retrieval System for Identifying Previously Fact-Checked Claims across Multiple Languages (Srijani Debnath, Pritam Pal and Dipankar Das)
- 12.00-12.20 LLM-based Embedders for Prior Case Retrieval (Damith Premasiri, Tharindu Ranasinghe and Ruslan Mitkov)

### **Parallel session 20: Deceptive Text Detection**

Room - Galata hall; Chair - Salima Lamsiya

- 11.20-11.40 Cross-Lingual Fact Verification: Analyzing LLM Performance Patterns across Languages (Hanna Shcharbakova, Tatiana Anikina, Natalia Skachkova and Josef van Genabith)
- 11.40-12.00 Detecting Deception in Disinformation across Languages: The Role of Linguistic Markers (Alba Perez-Montero, Silvia Gargova, Elena Lloret and Paloma Moreda Pozo)
- 12.00-12.20 Evaluating LLMs on Deceptive Text across Cultures (Katerina Papantoniou, Panagiotis Papadakos and Dimitris Plexousakis)

### **12.10-13.10 Student Research Workshop posters**

Chair: Petya Osenova

1. A Multi-Baseline Framework for Ranking Global Event Significance Using Google Trends and Large Language Models (Zenan Chen)
2. Investigating Hierarchical Structure in Multi-Label Document Classification (Artemis Dampa)
3. Large Language Models for Lexical Resource Enhancement: Multiple Hypernymy Resolution in WordNet (Dimitar Hristov)

4. NoCs: A Non-Compound-Stable Splitter for German Compounds (Carmen Schacht)
5. A Proposal for Evaluating the Linguistic Quality of Synthetic Spanish Corpora (Lucia Sevilla-Requena)
6. Visualization of LLM Annotated Documents (Teodor Todorov Valtchev and Nikolay Paev)

## 12.10-13.10 Poster session 2

Chair: Petya Osenova

### Posters:

1. QuARK: LLM-Based Domain-Specific Question Answering Using Retrieval Augmented Generation and Knowledge Graphs (Edward Burgin, Sourav Dutta and Mingxue Wang)
2. From Courtroom to Corpora: Building a Name Entity Corpus for Urdu Legal Texts (Adeel Zafar, Sohail Ashraf and Slawomir Nowaczyk)
3. Financial News as a Proxy of European Central Bank Interest Rate Adjustments (Davide Paris, Martina Menzio and Elisabetta Fersini)
4. LLM Compression: How Far Can We Go in Balancing Size and Performance? (Sahil Sk, Debashish Dhal, Sonal Khosla, Akash Dhaka, Shantipriya Parida, Sk Shahid, Sambit Shekhar, Dilip Prasad and Ondrej Bojar)
5. Chakoshi: A Customizable Guardrail for LLMs with a Focus on Japanese-Language Moderation (Kazuhiro Arai, Ryota Matsui, Kenji Miyama, Yudai Yamamoto, Ren Shibamiya, Kaito Sugimoto and Yoshimasa Iwase)
6. Instruction Finetuning to Attribute Language Stage, Dialect and Provenance Region to Historical Church Slavic Texts (Piroska Lendvai, Uwe Reichel, Anna Jouravel, Achim Rabus and Elena Renje)
7. C-SHAP: Collocation-Aware Explanations for Financial NLP (Martina Menzio, Elisabetta Fersini and Davide Paris)
8. Named Entity Recognition and Relation Extraction for Better Gut-Brain Interplay Understanding (Aleksis Ioannis Datseris, Mario Kuzmanov, Ivelina Nikolova-Koleva, Svetla Boytcheva and Dimitar Taskov)
9. Balancing the Scales: Addressing Gender Bias in Social Media Toxicity Detection (Beatriz Botella-Gil, Juan Pablo Consuegra-Ayala, Alba Bonet-Jover and Paloma Moreda-Pozo)
10. Authorship Verification Using Cloze Test with Large Language Models (Tomáš Foltýnek, Tomáš Kancko and Pavel Rychly)
11. The Challenge of Performing Ontology-driven Entity Extraction in Real-world Unstructured Textual Data from the Domain of Dementia (Sumaiya Suravee, Carsten Oliver Schmidt and Kristina Yordanova)

12. Detecting Fake News in the Era of Language Models (Muhammad Irfan Fikri Sabri, Hansi Hettiarachchi and Tharindu Ranasinghe)
13. Reddit-V: A Virality Prediction Dataset and Zero-Shot Evaluation with Large Language Models (Samir el-Amrany, Matthias R. Brust, Salima Lamsiyah and Pascal Bouvry)
14. Utilizing Large Language Models for Focused Conversational Assistants (Shruti Dhavalikar and Karthika Vijayan)
15. APIO: Automatic Prompt Induction and Optimization for Grammatical Error Correction and Text Simplification (Artem Chernodub, Aman Saini, Yejin Huh, Vivek Kulkarni and Vipul Raheja)
16. Pushing the (Generative) Envelope: Measuring the Effect of Prompt Technique and Temperature on the Generation of Model-based Systems Engineering Artifacts (Erin Smith Crabb, Cedric Bernard, Matthew Jones and Daniel Dakota)
17. Multi-Agent Reinforcement Learning for Interactive Code Debugging with Human Feedback and Memory (Anjana Krishnamoorthy, Kartik Ivatury and Benyamin Ahmadnia)
18. Differential Robustness in Transformer Language Models: Empirical Evaluation under Adversarial Text Attacks (Taniya Gidatkar, Oluwaseun Ajao and Matthew Shardlow)
19. MLDataForge: Accelerating Large-Scale Dataset Preprocessing and Access for Multimodal Foundation Model Training (Andrea Blasi Núñez, Lukas Paul Achatius Galke and Peter Schneider-Kamp)
20. Detecting Changes in Mental Health Status via Reddit Posts in Response to Global Negative Events (Zenan Chen, Judita Preiss and Peter A. Bath)
21. Modelling the Relative Contributions of Stylistic Features in Forensic Authorship Attribution (G. Çağatay Sat, John Blake and Evgeny Pyskhin)
22. Benchmarking Item Difficulty Classification in German Vocational Education and Training (Alonso Palomino and Benjamin Paassen)
23. Mitigating Bias in Text Classification via Prompt-Based Text Transformation (Charmaine Barker and Dimitar Kazakov)
24. Synthetic vs. Gold: The Role of LLM Generated Labels and Data in Cyberbullying Detection (Arefeh Kazemi, Sri Balaaji Natarajan Kalaivendan, Joachim Wagner, Hamza Qadeer, Kanishk Verma and Brian Davis)
25. From the Tractatus Logico-Philosophicus to Later Wittgenstein: A NLP-Based Comparative Analysis (Andreiana Mihail, Silviu-Florin Gheorghe, Andrei Fotea and Liviu P. Dinu)
26. Cyberbullying Detection via Aggression-Enhanced Prompting (Aisha Saeid, Anu Sabu, Girish Koushik, Ferrante Neri and Diptesh Kanojia)
27. Graph-based RAG for Low-Resource Aromanian–Romanian Translation (Laurentiu G. Ghetoiu and Sergiu Nisioi)
28. A Linguistically-informed Comparison between Multilingual BERT and Language-specific BERT Models: The Case of Differential Object Marking in Romanian (Maria Tepei and Jelke Bloem)
29. Can LLMs Disambiguate Grounded Language? The Case of PP Attachment (John Blackmore and Matthew Stone)
30. Optimism, Pessimism, and the Language between: Model Interpretability and Psycholinguistic Profiling (Stefana Arina Tabusca and Liviu P. Dinu)

- 31.F-LoRA-QA: Finetuning LLaMA Models with Low-Rank Adaptation for French Botanical Question Generation and Answering (Ayoub Nainia, Régine Vignes-Lebbe, Hajar Mousannif and Jihad Zahir)
- 32.Detecting Gender Stereotypical Language Using Model-agnostic and Model-specific Explanations (Manuela Nayantara Jeyaraj and Sarah Jane Delany)
- 33.Prompting Techniques for Reducing Social Bias in LLMs through System 1 and System 2 Cognitive Processes (Mahammed Kamruzzaman and Gene Louis Kim)
- 34.Reversing Causal Assumptions: Explainability in Online Sports Dialogues (Asteria Kaeberlein and Malihe Alikhani)
- 35.How LLMs Influence Perceived Bias in Journalism (Asteria Kaeberlein and Malihe Alikhani)
- 36.Exploring the Performance of Large Language Models for Event Detection and Extraction in the Domain of Health (Hristo Tanev, Nicolas Stefanovitch, Tomáš Harmatha and Diana F. Sousa)

**Remote poster presentations as pre-recorded videos:**

1. PersianSciQA: A New Dataset for Bridging the Language Gap in Scientific Question Answering (Safoura Aghadavoud Jolfaei, Azadeh Mohebi and Zahra Hemmat)
2. Task-Oriented Dialogue Systems through Function Calling (Tiziano Labruna, Giovanni Bonetta and Bernardo Magnini)
3. When to Retrieve: Teaching LLMs to Utilize Information Retrieval Effectively (Tiziano Labruna, Jon Ander Campos and Gorka Azkune)
4. The Erosion of LLM Signatures: Can We Still Distinguish Human and LLM-Generated Scientific Ideas after Iterative Paraphrasing? (Sadat Shahriar, Navid Ayoobi and Arjun Mukherjee)
5. A Question-Answering Based Framework/Metric for Evaluation of Newspaper Article Summarization (Vasanth Seemakurthy, Shashank Sundar, Siddharth Arvind, Siddhant Jagdish and Ashwini M. Joshi)
6. A Low-Resource Speech-Driven NLP Pipeline for Sinhala Dyslexia Assistance (Peshala Sandali Perera and Deshan Koshala Sumanathilaka)
7. PolyHope-M at RANLP2025 Subtask-1 Binary Hope Speech Detection: Spanish Language Classification Approach with Comprehensive Learning Using Transformer, and Traditional ML, and DL (Md. Julkar Naeen, Sourav Kumar Das, Sharun Akter Khushbu, Shahriar Sultan Ramit and Alaya Parven Alo)
8. Building a Clean Bartangi Language Corpus and Training Word Embeddings for Low-Resource Language Modeling (Warda Tariq, Victor Popov and Vasilii Gromov)
9. Harnessing Open-Source LLMs for Tender Named Entity Recognition (Asim Abbas, Venelin Kovatchev, Mark Lee, Niloofar Shanavas and Mubashir Ali)
10. Exploring the Usage of Knowledge Graphs in Identifying Human and LLM-Generated Fake Reviews (Ming Liu and Massimo Poesio)
11. Candidate Profile Summarization: A RAG Approach with Synthetic Data Generation for Tech Jobs (Anum Afzal, Ishwor Subedi and Florian Matthes)

13.10-15.10 **Lunch break**

**15.10-16.50 Parallel sessions 21-22 (regular papers: 20 min presentation; 5 min Q&A)**

**Parallel session 21: Understanding and Reasoning Skills**

Room - Chernomore hall; Chair - Ernesto Estebanel

15.10-15.35 Measuring How (Not Just Whether) VLMs Build Common Ground (Saki Imai, Mert Inan, Anthony B. Sicilia and Malihe Alikhani)

15.35-16.00 Can We Predict Innovation? Narrow Experts versus Competent Generalists (Amir Hazem and Motohashi Kazuyuki)

16.00-16.25 Enhancing Transformer-Based Rerankers with Synthetic Data and LLM-Based Supervision (Dimitar Peshevski, Kiril Blazhevski, Martin Popovski and Gjorgji Madjarov)

16.25-16.50

Check related pre-recorded presentations or papers in Proceedings:  
(short) Fast Thinking with Structured Prompts: Enabling LLM Reasoning without Chain-of-Thought Generation (Kirill Morozov, Liubov Chubarova and Irina Piontkovskaya)

**Parallel session 22: Simplification**

Room - Varna hall; Chair - Cengiz Acarturk

15.10-15.35 Towards CEFR-targeted Text Simplification for Question Adaptation (Luca Benedetto and Paula Buttery)

15.35-16.00 Simplifications are Absolutists: How Simplified Language Reduces Word Sense Awareness in LLM-Generated Definitions (Lukas Ellinger, Miriam Anschutz and Georg Groh)

16.00-16.25 Improving Estonian Text Simplification through Pretrained Language Models and Custom Datasets (Eduard Barbu, Meeri-Ly Muru and Sten Marcus Malva)

16.25-16.45 GPT-Based Lexical Simplification for Multi-Word Expressions Using Prompt Engineering (Sardar Khan Khayamkhani and Matthew Shardlow)

**15.10-16.50 Parallel session 23-24 (short papers: 15 min presentation; 5 min Q&A)**

**Parallel session 23: Multimodal Systems**

Room - Odessos Hall; Chair - Salmane Chafik

15.10-15.30 Fusion of Object-Centric and Linguistic Features for Domain-Adapted Multimodal Learning (Jordan Konstantinov Kralev)

15.30-15.50 Performance Gaps in Acted and Naturalistic Speech: Insights from Speech Emotion Recognition Strategies on Customer Service Calls (Lily Kawaoto, Hita Gupta, Ning Yu and Daniel Dakota)

15.50-16.10 Visual Priming Effect on Large-scale Vision Language Models (Daiki Yoshida, Haruki Sakajo, Kazuki Hayashi, Yusuke Sakai, Hidetaka Kamigaito, Katsuhiko Hayashi and Taro Watanabe)

Check related online presentations or papers in Proceedings: (short) Zero-shot OCR Accuracy of Low-Resourced Languages: A Comparative Analysis on Sinhala and Tamil (Nevidu Jayatilleke and Nisansa de Silva)

#### **Parallel session 24: Sentiment Analysis for Arabic**

Room - Galata hall; Chair - Saad Ezzini

15.10-15.30 Evaluating Large Language Models on Arabic Dialect Sentiment Analysis (Maram I. Alharbi, Saad Ezzini, Hansi Hettiarachchi, Tharindu Ranasinghe and Ruslan Mitkov)

15.30-15.50 An Annotated Dataset of Emotions and Hope Speech in English and Arabic (Wajdi Zaghouni and Md Rafiul Biswas)

15.50-16.10 An Annotated Corpus of Arabic Tweets for Hate Speech Analysis (Wajdi Zaghouni and Md Rafiul Biswas)

16.10-16.30 Evaluation of Pretrained and Instruction-Based Pretrained Models for Emotion Detection in Arabic Social Media Text (Md Rafiul Biswas, Shima Ibrahim, Mabrouka Bessghaier and Wajdi Zaghouni)

16.50-17.20 **Coffee break**

17.20-18.00 **Parallel sessions 25-28 (short papers: 15 min presentation; 5 min Q&A)**

#### **Parallel session 25: Multi-head attention, multi-aspect embedding and prompt engineering**

Room - Chernomore hall; Chair - Priya Dharshini Krishnaraj

17.20-17.40 A Deep Dive into Multi-Head Attention and Multi-Aspect Embedding (Maryam Teimouri, Jenna Kanerva and Filip Ginter)

17.40-18.00 Toponym Resolution: Will Prompt Engineering Change Expectations? (Isuri Anuradha, Deshan Koshala Sumanathilaka, Ruslan Mitkov and Paul Rayson)

#### **Parallel session 26: Question-Answering**

Room - Varna hall; Chair - Damith Premasiri

17.20-17.40 ASQ: Automatically Generating Question-Answer Pairs Using AMRs (Geetanjali Rakshit and Jeffrey Flanigan)

17.40-18.00 Advancing Active Learning with Ensemble Strategies (Naif Alatrush, Sultan Alsarra, Afraa Alshammari, Luay Abdeljaber, Niamat Zawad, Latifur Khan, Patrick T. Brandt, Javier Osorio and Vito D'Orazio)

### **Parallel session 27: Terminology**

Room - Odessos Hall; Chair - Marie Escribe

17.20-17.40 The Evaluation of Medical Terms Complexity Using Lexical Features and Large Language Models (Liliya Makhmutova, Giancarlo Dondoni Salton, Fernando Perez-Tellez and Robert J. Ross)

Check related online presentations or papers in Proceedings: (short) Evaluating Large Language Models on Multiword Expressions in Multilingual and Code-Switched Contexts (Frances Adriana Laureano De Leon, Asim Abbas, Harish Tayyar Madabushi and Mark Lee)

### **Parallel session 28: Advances in Language Encoding**

Room - Galata hall; Chair - Tharindu Ranasinghe

17.20-17.40 ExPe: Exact Positional Encodings for Generative Transformer Models with Extrapolating Capabilities (Aleksis Ioannis Datseris, Sylvia Vassileva, Ivan K. Koychev and Svetla Boytcheva)

17.40-18.00 Domain Knowledge Distillation for Multilingual Sentence Encoders in Cross-lingual Sentence Similarity Estimation (Risa Kondo, Hiroki Yamauchi, Tomoyuki Kajiwara, Marie Katsurai and Takashi Ninomiya)

18.00-19.00 **Cambridge University Press (CUP)+Journal 'Natural Language Processing' (JNLP) presentation**

Host: Ruslan Mitkov

19.00-19.15 **Official closing of the main conference**

19.30 **Welcome Workshops and Shared Tasks Cocktail Reception sponsored by Cambridge University Press**