



**Bulgarian Academy of Sciences
Institute of Information and
Communication Technologies
Department of Artificial Intelligence
and Language Technologies**

PROGRAMME

of the 14th International Conference on

Recent Advances in Natural Language Processing

RANLP-2023



4 - 6 September, 2023

CHERNO MORE Hotel, Varna, Bulgaria



Grant КП-06-МНФ/3, 19.05.2023



ontotext

IRIS.AI

senso

RANLP 2023 Main Conference Programme with chairs and lecturing halls

Day 1 (4 September 2023)

9.00-9.15 Opening (Cherno more hall)

9.15-10.15 Keynote speech 1: Eduard Hovy

The complementarity of neural and symbolic approaches to NLP

Cherno more hall (40-45 min presentation; 10 min Q & A)

Introduced by **Ruslan Mitkov**

10.45-1.15 Parallel sessions 1-4 (regular papers: 20 min presentation; 5 min Q&A)

Parallel Session 1: Language Resources and Multilinguality

Session Chair: Erhard Hinrichs, Cherno more hall

10.45-11.10 Studying Common Ground Instantiation Using Audio, Video and Brain Behaviours: the BrainKT corpus (Eliot Maës, Thierry Legou, Leonor Becerra and Philippe Blache)

11.10-11.35 Transformer-Based Language Models for Bulgarian (Iva Marinova, Kiril Simov and Petya Osenova)

11.35-12.00 Developing a Multilingual Corpus of Wikipedia Biographies (Hannah Devinney, Anton Eklund, Igor Ryazanov and Jingwen Cai)

12.00-12.25 WIKITIDE: A Wikipedia-based Timestamped Definition Pairs Dataset (Hsuvas Borkakoty and Luis Espinosa Anke)

12.25-12.50 Automatic Extraction of the Romanian Academic Word List: Data and Methods (Ana-Maria Bucur, Andreea Dincă, Madalina Chitez and Roxana Rogobete)

12.50-13.15 Pretraining Language- and Domain-Specific BERT on Automatically Translated Text (Tatsuya Ishigaki, Yui Uehara, Goran Topić and Hiroya Takamura)

Parallel session 2: Language Models: Training, Adaptation and Evaluation

Session Chair: Damith Premasiri, Varna hall

10.45-11.10 Bridging the Gap between Subword and Character Segmentation in Pretrained Language Models (Shun Kiyono, Sho Takase, Shengzhe Li and Toshinori Sato)

11.10-11.35 Forming Trees with Treeformers (Nilay Patel and Jeffrey Flanigan)

11.35-12.00 Modeling Easiness for Training Transformers with Curriculum Learning (Leonardo Ranaldi, Giulia Pucci and Fabio Massimo Zanzotto)

12.00-12.25 Towards a Consensus Taxonomy for Annotating Errors in Automatically Generated Text (Rudali Huidrom and Anya Belz)

12.25-12.50 Evaluating of Large Language Models in Relationship Extraction from Unstructured Data: Empirical Study from Holocaust Testimonies (Isuri Anuradha, Le An Ha, Ruslan Mitkov and Vinita Nahar)

12.50-13.15 Lessons Learnt from Linear Text Segmentation: a Fair Comparison of Architectural and Sentence Encoding Strategies for Successful Segmentation (Iacopo Ghinassi, Lin Wang, Chris Newell and Matthew Purver)

Parallel session 3: Fake, Bias and Hate Speech Detection

Session Chair: Sandra Kübler, Galata hall

10.45-11.10 Looking for Traces of Textual Deepfakes in Bulgarian on Social Media (Irina Temnikova, Iva Marinova, Silvia Gargova, Ruslana Margova and Ivan Koychev)

11.10-11.35 Multi-task Ensemble Learning for Fake Reviews Detection and Helpfulness Prediction: A Novel Approach (Alimuddin Melleng, Anna Jurek-Loughrey and Deepak P)

11.35-12.00 From Fake to Hyperpartisan News Detection Using Domain Adaptation (Răzvan-Alexandru Smădu, Sebastian-Vasile Echim, Dumitru-Clementin Cercel, Iuliana Marin and Florin Pop)

12.00-12.25 Bipol: Multi-axes Evaluation of Bias with Explainability in Benchmark Datasets (Tosin Adewumi, Isabella Södergren, Lama Alkhaled, sana al-azzawi, Foteini Simistira Liwicki and Marcus Liwicki)

12.25-12.50 Exploring Amharic Hate Speech Data Collection and Classification Approaches (Abinew Ali Ayele, Seid Muhie Yimam, Tadesse Destaw Belay, Tesfa Asfaw and Chris Biemann)

12.50-13.15 Prompt-Based Approach for Czech Sentiment Analysis (Jakub Šmíd and Pavel Přibáň)

Parallel session 4: Sentence-level Representation and Analysis

Session chair: Ivelina Nikolova, Odessos Hall

10.45-11.10 Efficient Domain Adaptation of Sentence Embeddings Using Adapters (Tim Schopf, Dennis Schneider and Florian Matthes)

11.10-11.35 AspectCSE: Sentence Embeddings for Aspect-based Semantic Textual Similarity using Contrastive Learning and Structured Knowledge (Tim Schopf, Emanuel Gerber, Malte Ostendorff and Florian Matthes)

11.35-12.00 Comparative Analysis of Anomaly Detection Algorithms in Text Data (Yizhou Xu, Kata Gábor, Jérôme Milleret and Frédérique Segond)

12.00-12.25 Simultaneous Interpreting as a Noisy Channel: How Much Information Gets Through (Maria Kunilovskaya, Heike Przybyl, Ekaterina Lapshinova-Koltunski and Elke Teich)

12.25-12.50 Explainable Event Detection with Event Trigger Identification as Rationale Extraction (Hansi Hettiarachchi and Tharindu Ranasinghe)

12.50-13.15 Word Sense Disambiguation for Automatic Translation of Medical Dialogues into Pictographs (Magali Norré, Rémi Cardon, Vincent Vandeghinste and Thomas François)

'Remote regular papers': available on the conference website to view on 4, 5 and 6 September 2023 only:

1. 3D-EX: A Unified Dataset of Definitions and Dictionary Examples (Fatemah Almeman, Hadi Sheikhi and Luis Espinosa Anke)
2. Data Augmentation for Fake News Detection by Combining Seq2seq and NLI (Anna Glazkova)

1.15-2.45pm **Lunch break**

2.45-3.45pm **Keynote 2:** Tharindu Ranasinghe
Responsible Machine Learning: Challenges and Way Forward?
Cherno more hall (40-45 min presentation; 10 min Q & A)

Introduced by **Ruslan Mitkov**

3.45-4.30pm **Diamond sponsor presentation** (Ontotext), Cherno more hall

Introduced by **Galia Angelova**

4.30-5.00pm **Coffee break**

5.00-5.30pm **Gold sponsor presentation** (Iris.ai), Cherno more hall

Introduced by **Ruslan Mitkov**

5.30-6.30pm **Demo session**, Cherno more hall

Session chair: **Galia Angelova**

NoHateBrazil: A Brazilian Portuguese Text Offensiveness Analysis System (Francielle Vargas, Isabelle Carvalho, Wolfgang Schmeisser-Nieto, Fabrício Benevenuto and Thiago Pardo)

!Translate: When You Cannot Cook Up a Translation, Explain (Federico Garcea, Margherita Martinelli, Maja Milicević Petrović and Alberto Barrón-Cedeño)

T2KG: Transforming Multimodal Document to Knowledge Graph (Santiago Galiano, Rafael Muñoz, Yoan Gutiérrez, Andrés Montoyo, Jose Ignacio Abreu and Luis Alfonso Ureña)

Vocab-Expander: A System for Creating Domain-Specific Vocabularies Based on Word Embeddings (Michael Faerber and Nicholas Popovic)

6.00-7.00pm **Poster session 1**, Varna hall

Session Chair Tharindu Ransinghe

22 posters: Varieties of topics covering

- text classification for approaching a range of problems from hate speech and propaganda detection in mass media to applications in biomedical domain, education and business, including in multilingual context;
 - question-answering and dialogue systems;
 - topic modelling
1. Multilingual Racial Hate Speech Detection Using Transfer Learning (Abinew Ali Ayele, Skadi Dinter, Seid Muhie Yimam and Chris Biemann)
 2. Propaganda Detection in Russian Telegram Posts in the Scope of the Russian Invasion of Ukraine (Natalia Vanetik, Marina Litvak, Egor Reviakin and Margarita Tiamanova)
 3. Clinical text classification to SNOMED CT codes using transformers trained on linked open medical ontologies (Anton Hristov, Petar Ivanov, Anna Aksenova, Tsvetan Asamov, Pavlin Gyurov, Todor Primov and Svetla Boytcheva)
 4. Exploring Unsupervised Semantic Similarity Methods for Claim Verification in Health Care News Articles (Vishwani Gupta, Astrid Viciano, Holger Wormer and Najmehsadat Mousavinezhad)
 5. Context Aware Module Selection in Modular Dialog Systems (Jan Nehring, René Marcel Berk and Stefan Hillmann)
 6. Source Code Plagiarism Detection with Pre-Trained Model Embeddings and Automated Machine Learning (Fahad Ebrahim and Mike Joy)
 7. BanglaBait: Semi-Supervised Adversarial Approach for Clickbait Detection on Bangla Clickbait Dataset (Md. Motahar Mahtab, Monirul Haque, Mehedi Hasan and Farig Sadeque)
 8. Classification-Aware Neural Topic Model Combined With Interpretable Analysis - For Conflict Classification (Tianyu Liang, Yida Mu, Soonho Kim, Darline Kuate, Julie Lang, Rob Vos and Xingyi Song)
 9. 'ChemXtract' A System for Extraction of Chemical Events from Patent Documents (Pattabhi RK Rao and Sobha Lalitha Devi)
 10. Challenges of GPT-3-based Conversational Agents for Healthcare (Fabian Lechner, Allison Lahnala, Charles Welch and Lucie Flek)
 11. Three Approaches to Client Email Topic Classification (Branislava Šandrih Todorović, Katarina Josipović and Jurij Kodre)
 12. Multimodal Learning for Accurate Visual Question Answering: An Attention-based Approach (Jishnu Bhardwaj, Anurag Balakrishnan, Satyam Pathak, Ishan Unnarkar, Aniruddha Gawande and Benyamin Ahmadnia)
 13. Uncertainty Quantification of Text Classification in a Multi-Label Setting for Risk-Sensitive Systems (Jinha Hwang, Carol Gudumotu and Benyamin Ahmadnia)
 14. A Review in Knowledge Extraction from Knowledge Bases (Fabio Yanez, Andrés Montoyo, Yoan Gutierrez, Rafael Muñoz and Armando Suarez)

15. Exploring the Landscape of Natural Language Processing Research (Tim Schopf, Karim Arabi and Florian Matthes)
16. Reading Between the Lines: Information Extraction from Industry Requirements (Ole Magnus Holter and Basil Ell)
17. Multilingual Continual Learning Approaches for Text Classification (Karan Praharaj and Irina Matveeva)
18. Auto-Encoding Questions with Retrieval Augmented Decoding for Unsupervised Passage Retrieval and Zero-shot Question Generation (Stalin Varanasi, Muhammad Umer Tariq Butt and Guenter Neumann)
19. On the generalization of projection-based gender debiasing in word embedding (Elisabetta Fersini, ANTONIO CANDELIERI and Lorenzo Pastore)
20. A Review of Research-based Automatic Text Simplification Tools (Isabel Espinosa-Zaragoza, José Abreu-Salas, Elena Lloret Pastor, Paloma Moreda Pozo and Manuel Palomar Sanz)
21. huPWKP: A Hungarian Text Simplification Corpus (Noémi Prótár and Dávid Márk Nemeskey)
22. Evaluating Unsupervised Hierarchical Topic Models Using a Labeled Dataset (Judicael Poumay and Ashwin Ittoo)
23. A Computational Analysis of the Voices of Shakespeare's Characters (Liviu P. Dinu and Ana Sabina Uban)

'Remote posters': available on the conference website to view on 4, 5 and 6 September 2023 only

1. Kāraka-Based Answer Retrieval for Question Answering in Indic Languages (Devika Verma, Ramprasad S. Joshi, Aiman A. Shivani and Rohan D. Gupta)
2. HAQA and QUQA: Constructing two Arabic Question-Answering Corpora for the Quran and Hadith (Sarah Alnefaie, Eric Atwell and Mohammad Ammar Alsalka)

8.30pm-2.00am **Gala dinner**

Day 2 (5 September 2023)

9.30-10.30 Keynote 3: Sandra Kübler
Hate Speech Detection: What have we learned?
Cherno more hall (40-45 min presentation; 10 min Q&A)

Introduced by Galia Angelova

10.30-11.00 Coffee break

11-12.40 Parallel sessions 5-8 (short papers: 15 min presentation; 5 min Q&A)

Parallel session 5: Multimodal Studies

Session Chair: Richard Rapp Odessos Hall

11.00-11.20 Sign Language Recognition and Translation: A Multi-Modal Approach Using Computer Vision and Natural Language Processing (Jacky Li, Jaren Gerdes, James Gojit, Austin Tao, Samyak Katke, Kate Nguyen and Benyamin Ahmadnia)

11.20-11.40 Unimodal Intermediate Training for Multimodal Meme Sentiment Classification (Muzhaffar Hazman, Susan McKeever and Josephine Griffith)

11.40-12.00 Was That a Question? Automatic Classification of Discourse Meaning in Spanish (Santiago Arróniz and Sandra Kübler)

12.00-12.20 MQDD: Pre-training of Multimodal Question Duplicity Detection for Software Engineering Domain (Jan Pasek, Jakub Sido, Miloslav Konopik and Ondrej Prazak)

12.20-12.40 From Stigma to Support: A Parallel Monolingual Corpus and NLP Approach for Neutralizing Mental Illness Bias (Mason Choey)

Parallel session 6: NLP for Social Sciences and Literary Studies

Session Chair: Ivan Koychev Galata hall

11.00-11.20 ConflIBERT-Arabic: A Pre-trained Arabic Language Model for Politics, Conflicts and Violence (Sultan Alsarra, Luay Abdeljaber, Wooseong Yang, Niamat Zawad, Latifur Khan, Patrick Brandt, Javier Osorio and Vito D'Orazio)

11.20-11.40 Detecting Text Formality: A Study of Text Classification Approaches (Daryna Dementieva, Nikolay Babakov and Alexander Panchenko)

11.40-12.00 Taxonomy-Based Automation of Prior Approval Using Clinical Guidelines (Saranya Krishnamoorthy, Ayush Singh and John E. Ortega)

12.00-12.20 Using Wikidata for Enhancing Compositionality in Pretrained Language Models (Meriem Beloucif, Mihir Bansal and Chris Biemann)

12.20-12.40 Deep Learning Methods for Identification of Multiword Flower and Plant Names (Damith Premasiri, Amal Haddad Haddad, Tharindu Ranasinghe and Ruslan Mitkov)

Parallel session 7: Machine Translation

Chair: Maria Kunilovskaya, Varna hall

11.00-11.20 Improving Translation Quality for Low-Resource Inuktitut with Various Preprocessing Techniques (Mathias Hans Erik Stenlund, Mathilde Nanni, Micaella Bruton and Meriem Beloucif)

11.20-11.40 An Evaluation of Source Factors in Concatenation-based Context-aware Neural Machine Translation (Harritxu Gete and Thierry Etchegoyhen)

11.40-12.00 Measuring Spurious Correlation in Classification: "Clever Hans" in Translationese (Angana Borah, Daria Pylypenko, Cristina España-Bonet and Josef van Genabith)

12.00-12.20 TreeSwap: Data Augmentation for Machine Translation via Dependency Subtree Swapping (Attila Nagy, Dorina Lakatos, Botond Barta and Judit Ács)

12.20-12.40 Cross-lingual Classification of Crisis-related Tweets Using Machine Translation (Shareefa Al Amer, Mark Lee and Phillip Smith)

Parallel session 8: Opinion-mining and Sentiment Analysis

Session Chair: Alistair Plum, Chernomore hall

11.00-11.20 Measuring Gender Bias in Natural Language Processing: Incorporating Gender-Neutral Linguistic Forms for Non-Binary Gender Identities in Abusive Speech Detection (Nasim Sobhani, Kinshuk Sengupta and Sarah Jane Delany)

11.20-11.40 Beyond Information: Is ChatGPT Empathetic Enough? (Ahmed Belkhir and Fatiha Sadat)

11.40-12.00 Impact of Emojis on Automatic Analysis of Individual Emotion Categories (Ratchakrit Arreerard and Scott Piao)

12.00-12.20 SSSD: leveraging pre-trained models and semantic search for semi-supervised stance detection (André Mediate de Sousa and Karin Becker)

12.20-12.40 Mapping Explicit and Implicit Discourse Relations between the RST-DT and the PDTB 3.0 (Nelson Filipe Costa, Nadia Sheikh and Leila Kosseim)

‘Remote short papers’: available on the conference website to view on 4, 5 and 6 September 2023 only:

1. AlphaMWE-Arabic: Arabic Edition of Multilingual Parallel Corpora with Multiword Expression Annotations (Najet Hadj Mohamed, Malak Rassem, Lifeng Han and Goran Nenadic)
2. A Research-Based Guide for the Creation and Deployment of a Low-Resource Machine Translation System (John E. Ortega and Kenneth Church)

12.40-2.00 pm **Lunch break**

2.00-3.00 **Keynote 4:** Lucas Beyer

The convergence of vision and language

Chernomore hall (40-45 min presentation; 10 min Q & A)

Introduced by Ruslan Mitkov

3.15pm-9.30pm **Excursion**

Day 3 (6 September 2023)

9.00-10.00 **Keynote 5:** Isabelle Augenstein
Quantifying gender biases towards entities
Cherno more hall (40-45 min presentation; 10 min Q & A)

Introduced by **Ruslan Mitkov**

10.00-10.30 **Coffee break**

11.00-1.05 **Parallel session 9: Language Models for Text Understanding and Production**

(regular papers: 20 min presentation; 5 min Q&A)

Session Chair: Kiril Simov, Varna hall

11.00-11.25 Does the "most sinfully decadent cake ever" taste good? Answering Yes/No Questions from Figurative Contexts (Geetanjali Rakshit and Jeffrey Flanigan)

11.25-11.50 Bigfoot in Big Tech: Detecting Out of Domain Conspiracy Theories (Matthew Fort, Zuoyu Tian, Elizabeth Gabel, Nina Georgiades, Noah Sauer, Daniel Dakota and Sandra Kübler)

11.50-12.15 Exploring Abstractive Text Summarisation for Podcasts: A Comparative Study of BART and T5 Models (Parth Saxena and Mahmoud El-Haj)

12.15-12.40 Classifying COVID-19 Vaccine Narratives (Yue Li, Carolina Scarton, Xingyi Song and Kalina Bontcheva)

12.40-1.05 Party Extraction from Legal Contract Using Contextualized Span Representations of Parties (Sanjeevan Sivapiran, Charangan Vasantharajan and Uthayasanker Thayasivam)

10.30-1.30 **Parallel sessions 10-12** (short papers: 15 min presentation; 5 min Q&A)

Parallel session 10: LM under Scrutiny: Explanation, Evaluation, Fine-tuning Ideas and Applications

Session Chair: Isuri Anuradha, Odessos hall

10.30-10.50 The Dark Side of the Language: Pre-trained Transformers in the DarkNet (Leonardo Ranaldi, Aria Nourbakhsh, Elena Sofia Ruzzetti, Arianna Patrizi, Dario Onorati, Michele Mastromattei, Francesca Fallucchi and Fabio Massimo Zanzotto)

10.50-11.10 PreCog: Exploring the Relation between Memorization and Performance in Pre-trained Language Models (Leonardo Ranaldi, Elena Sofia Ruzzetti and Fabio Massimo Zanzotto)

11.10-11.30 Evaluating Generative Models for Graph-to-Text Generation (Shuzhou Yuan and Michael Faerber)

11.30-11.50 Can Model Fusing Help Transformers in Long Document Classification? An Empirical Study (Damith Premasiri, Tharindu Ranasinghe and Ruslan Mitkov)

11.50-12.10 Microsyntactic Unit Detection using Word Embedding Models: Experiments on Slavic Languages (Iuliia Zaitova, Irina Stenger and Tania Avgustinova)

12.10-12.30 BERTabaporu: assessing a genre-specific language model for Portuguese NLP (Pablo Botton Costa, Matheus Camasmie Pavan, Wesley Ramos Santos, Samuel Caetano Silva and Ivandré Paraboni)

12.30-12.50 Mind the User! Measures to More Accurately Evaluate the Practical Value of Active Learning Strategies (Julia Romberg)

12.50-1.10 Publish or Hold? Automatic Comment Moderation in Luxembourgish News Articles (Tharindu Ranasinghe, Alistair Plum, Christoph Purschke and Marcos Zampieri)

1.10-1.30 Poetry Generation Combining Poetry Theme Labels Representations (Yingyu Yan, Dongzhen Wen, Liang Yang, Dongyu Zhang and Hongfei LIN)

Parallel session 11: LM for Detecting Fakes, Misinformation, Deception, Bias, Intent

Session Chair: Petya Osenova, Chernomore hall

10.30-10.50 Categorising Fine-to-Coarse Grained Misinformation: An Empirical Study of the COVID-19 Infodemic (Ye Jiang, Xingyi Song, Carolina Scarton, Iknor Singh, Ahmet Aker and Kalina Bontcheva)

10.50-11.10 Data Fusion for Better Fake Reviews Detection (Alimuddin Melleng, Anna Jurek-Loughrey and Deepak P)

11.10-11.30 Socially Responsible Hate Speech Detection: Can Classifiers Reflect Social Stereotypes? (Francielle Vargas, Isabelle Carvalho, Ali Hürriyetoglu, Thiago Pardo and Fabrício Benevenuto)

11.30-11.50 Noisy Self-Training with Data Augmentations for Offensive and Hate Speech Detection Tasks (João Leite, Carolina Scarton and Diego Silva)

11.50-12.10 Predicting Sentence-Level Factuality of News and Bias of Media Outlets (Francielle Vargas, Kokil Jaidka, Thiago Pardo and Fabrício Benevenuto)

12.10-12.30 On the Identification and Forecasting of Hate Speech in Inceldom (Paolo Gajo, Arianna Muti, Katerina Korre, Silvia Bernardini and Alberto Barrón-Cedeño)

12.30-12.50 Deep Learning Approaches to Detecting Safeguarding Concerns in Schoolchildren's Online Conversations (Emma Franklin and Tharindu Ranasinghe)

12.50-1.10 Enriched Pre-trained Transformers for Joint Slot Filling and Intent Detection (Momchil Hardalov, Ivan Koychev and Preslav Nakov)

1.10-1.30 Cross-Lingual Speaker Identification for Indian Languages (Amaan Rizvi, Anupam Jamatia, Dwijen Rudrapal, Kunal Chakma and Björn Gambäck)

Parallel session 12: Simplification, Summarisation, Generation, Classification and Resources

Session Chair: Galia Angelova, Galata hall

10.30-10.50 Systematic TextRank Optimization in Extractive Summarization (Morris Zieve, Anthony Gregor, Frederik Juul Stokbaek, Hunter Lewis, Ellis Marie Mendoza and Benyamin Ahmadnia)

10.50-11.10 Non-Parametric Memory Guidance for Multi-Document Summarization (Florian Baudand Alex Aussem)

11.10-11.30 BB25HLegalSum: Leveraging BM25 and BERT-based clustering for the summarization of

legal documents (Leonardo Bonalume de Andrade and Karin Becker)

11.30-11.50 Advancing Topical Text Classification: A Novel Distance-Based Method with Contextual Embeddings (Andriy Kosar, Guy De Pauw and Walter Daelemans)

11.50-12.10 HTMOT : Hierarchical Topic Modelling Over Time (Judicael POUMAY and Ashwin Ittoo)

12.10-12.30 Classification of US Supreme Court Cases using BERT-Based Techniques (ShubhamVatsal, Adam Meyers and John E. Ortega)

12.30-12.50 Designing the LECOR Learner Corpus for Romanian (Ana Maria Barbu, Elena Irimia, Carmen Mîrzea Vasile and Vasile Păiș)

12.50-1.10 Lexicon-driven automatic sentence generation for the skills section in a job posting (Vera Aleksic, Mona Brems, Anna Mathes and Theresa Bertele)

1.10-1.30 A Computationally-Light Lexical Simplifier for Spanish (Sanja Stajner, Daniel Ibanez and Horacio Saggion)

'Remote short papers': available on the conference website to view on 4, 5 and 6 September 2023 only:

1. Tackling the Myriads of Collusion Scams on YouTube Comments of Cryptocurrency Videos (Sadat Shahriar and Arjun Mukherjee)
2. Data Augmentation for Fake Reviews Detection (Ming Liu and Massimo Poesio)
3. Exploring Deceptive Domain Transfer Strategies: Mitigating the Differences among Deceptive Domains (Sadat Shahriar, Arjun Mukherjee and Omprakash Gnawali)
4. Student's t-Distribution: On Measuring the Inter-Rater Reliability When the Observations are Scarce (Serge Gladkoff, Lifeng Han and Goran Nenadic)
5. Coherent Story Generation with Structured Knowledge (Congda Ma, Kotaro Funakoshi, Kiyooki Shirai and Manabu Okumura)
6. A Practical Survey on Zero-Shot Prompt Design for In-Context Learning (Yinheng Li)
7. Automatically Generating Hindi Wikipedia Pages using Wikidata as a Knowledge Graph: A Domain-Specific Template Sentences Approach (Aditya Agarwal and Radhika Mamidi)
8. Performance analysis of Arabic pre-trained models on named entity recognition task (Abdelhalim Hafedh Dahou, Mohamed Amine Cheragui and Ahmed Abdelali)
9. Taxonomy-Based Automation of Prior Approval using Clinical Guidelines (Saranya Krishnamoorthy and Ayush Singh)
10. Topic Modeling Using Community Detection on a Word Association Graph (Mahfuzur Rahman Chowdhury, Intesur Ahmed, Farig Sadeque and Muhammad Yanhaona)

1.30-3.00pm **Lunch break**

3.00-4.00 **Keynote 6**: Efstathios Stamatatos

Exploring Advances and Challenges in Authorship Verification

Cherno more hall (40-45 min presentation; 10 min Q & A)

Introduced by Galia Angelova

4.00-4.30 **Coffee break**

4.30-4.45 **Silver sponsor presentation** (Senso), Chernomore hall

Introduced by Ruslan Mitkov

4.45-5.45 **Poster session 2 incorporating the student workshop**

Chair: Karin Becker, Varna hall

12 posters on NER, Sentiment Analysis, Assessment in Education and Machine Translation

1. Evaluating Data Augmentation for Medication Identification in Clinical Notes (Jordan Koontz, Maite Oronoz and Alicia Pérez)
2. Medical Concept Mention Identification in Social Media Posts using a Small Number of Sample References (Vasudevan Nedumpozhimana, Sneha Rautmare, Meegan Gower, Nishtha Jain, Maja Popović, Patricia Buffini and John Kelleher)
3. Comparison of Multilingual Entity Linking Approaches (Ivelina Bozhinova and Andrey Tagarev)
4. Identifying Semantic Argument Types in Predication and Copredication Contexts: A Zero-Shot Cross-Lingual Approach (Deniz Ekin Yavas, Laura Kallmeyer, Rainer Osswald, Elisabetta Jezek, Marta Ricchiardi and Long Chen)
5. Exploring Techniques to Detect and Mitigate Non-Inclusive Language Bias in Marketing Communications using a Dictionary-Based Approach (Bharathi Raja Chakravarthi, Prasanna Kumar Kumaresan, Rahul Ponnusamy, John P. McCrae, Michaela Comerford, Jay Megaro, Deniz Keles and Last Feremenga)
6. Human Value Detection from Bilingual Sensory Product Reviews (Boyu Niu, Céline Manetta and Frédérique Segond)
7. Stance Prediction from Multimodal Social Media Data (Lais Carraro Leme Cavalheiro, Matheus Camasmie Pavan and Ivandré Paraboni)
8. Improving Aspect-Based Sentiment with End-to-End Semantic Role Labeling Model (Pavel Přibáň and Ondřej Pražák)
9. Dimensions of Quality: Contrasting Stylistic vs. Semantic Features for Modelling Literary Quality in 9,000 Novels (Pascale Moreira and Yuri Bizzoni)
10. Discourse Analysis of Argumentative Essays of English Learners based on CEFR Level (Blaise Hanel and Leila Kosseim)
11. Hindi to Dravidian Language Neural Machine Translation Systems (Vijay Sundar Ram and Sobha Lalitha Devi)

Posters from the student workshop

'Remote posters': available on the conference website to view on 4, 5 and 6 September 2023 only
WordNet: Problems in Translating Hindi Synsets into Bhojpuri (Imran Ali and Praveen Gatla)

1. Generative Models For Indic Languages: Evaluating Content Generation Capabilities (Savita

Bhat, Vasudeva Varma and Niranjana Pedaneekar)

2. Event Annotation and Detection in Kannada-English Code-Mixed Social Media Data (Sumukh S, Abhinav Appidi and Manish Shrivastava)
3. Comparative Analysis of Named Entity Recognition in the Dungeons and Dragons Domain (Gayashan Weerasundara and Nisansa de Silva)
4. Automatic Assessment of Spoken English Proficiency Based on Multimodal & Multitask Transformers (Kamel Nebhi and György Szaszák)

5.45-6.00 **Closing**, Chernomore hall

6.00-7.30 **JNLE promotional event**, Chernomore hall

8.30-11.00 **Reception sponsored by Cambridge University Press**, Varna hall