

# **International Conference on Recent Advances in Natural Language Processing (RANLP 2023)**

Large Language Models for Natural  
Language Processing

Varna, Bulgaria  
4 – 6 September 2023

Volume 1 of 2

**Editors:**

**Galia Angelova  
Maria Kunilovskaya  
Ruslan Mitkov**

ISBN: 978-1-7138-8240-4

**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© (2023) by the Association for Computational Linguistics  
All rights reserved.

Printed with permission by Curran Associates, Inc. (2024)

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](mailto: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](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>Bipol: Multi-Axes Evaluation of Bias with Explainability in Benchmark Datasets</i> Tosin Adewumi, Isabella Södergren, Lama Alkhaled, sana al-azzawi, Foteini Simistira Liwicki and Marcus Liwicki .....	1
<i>Automatically Generating Hindi Wikipedia Pages Using Wikidata as a Knowledge Graph: A Domain-Specific Template Sentences Approach</i> Aditya Agarwal and Radhika Mamidi .....	11
<i>Cross-lingual Classification of Crisis-related Tweets Using Machine Translation</i> Shareefa Al Amer, Mark Lee and Phillip Smith .....	22
<i>Lexicon-Driven Automatic Sentence Generation for the Skills Section in a Job Posting</i> Vera Aleksic, Mona Brems, Anna Mathes and Theresa Bertele .....	32
<i>Multilingual Racial Hate Speech Detection Using Transfer Learning</i> Abinew Ali Ayele, Skadi Dinter, Seid Muhie Yimam and Chris Biemann .....	41
<i>Exploring Amharic Hate Speech Data Collection and Classification Approaches</i> Abinew Ali Ayele, Seid Muhie Yimam, Tadesse Destaw Belay, Tesfa Asfaw and Chris Biemann	49
<i>Bhojpuri WordNet: Problems in Translating Hindi Synsets into Bhojpuri</i> Imran Ali and Praveen Gatla .....	60
<i>3D-EX: A Unified Dataset of Definitions and Dictionary Examples</i> Fatemah Almeman, Hadi Sheikhi and Luis Espinosa Anke .....	69
<i>Are You Not moved? Incorporating Sensorimotor Knowledge to Improve Metaphor Detection</i> Ghadi Alnafesah, Phillip Smith and Mark Lee .....	80
<i>HAQA and QUQA: Constructing Two Arabic Question-Answering Corpora for the Quran and Hadith</i> Sarah Alnefaie, Eric Atwell and Mohammad Ammar Alsalka .....	90
<i>ConflBERT-Arabic: A Pre-trained Arabic Language Model for Politics, Conflicts and Violence</i> Sultan Alsarra, Luay Abdeljaber, Wooseong Yang, Niamat Zawad, Latifur Khan, Patrick Brandt, Javier Osorio and Vito D’Orazio .....	98
<i>A Review in Knowledge Extraction from Knowledge Bases</i> Fabio Yanez, Andrés Montoyo, Yoan Gutierrez, Rafael Muñoz and Armando Suarez .....	109
<i>Evaluating of Large Language Models in Relationship Extraction from Unstructured Data: Empirical Study from Holocaust Testimonies</i> Isuri Anuradha, Le An Ha, Ruslan Mitkov and Vinita Nahar .....	117
<i>Impact of Emojis on Automatic Analysis of Individual Emotion Categories</i> Ratchakrit Arreerard and Scott Piao .....	124
<i>Was That a Question? Automatic Classification of Discourse Meaning in Spanish</i> Santiago Arróniz and Sandra Kübler .....	132
<i>Designing the LECOR Learner Corpus for Romanian</i> Ana Maria Barbu, Elena Irimia, Carmen Mîrzea Vasile and Vasile Păiș .....	143

<i>Non-Parametric Memory Guidance for Multi-Document Summarization</i> Florian Baud and Alex Aussem .....	153
<i>Beyond Information: Is ChatGPT Empathetic Enough?</i> Ahmed Belkhir and Fatiha Sadat .....	159
<i>Using Wikidata for Enhancing Compositionality in Pretrained Language Models</i> Meriem Beloucif, Mihir Bansal and Chris Biemann .....	170
<i>Multimodal Learning for Accurate Visual Question Answering: An Attention-Based Approach</i> Jishnu Bhardwaj, Anurag Balakrishnan, Satyam Pathak, Ishan Unnarkar, Aniruddha Gawande and Benyamin Ahmadnia .....	179
<i>Generative Models For Indic Languages: Evaluating Content Generation Capabilities</i> Savita Bhat, Vasudeva Varma and Niranjana Pedanekar .....	187
<i>Measuring Spurious Correlation in Classification: "Clever Hans" in Translationese</i> Angana Borah, Daria Pylypenko, Cristina España-Bonet and Josef van Genabith .....	196
<i>WIKITIDE: A Wikipedia-Based Timestamped Definition Pairs Dataset</i> Hsuvas Borkakoty and Luis Espinosa Anke .....	207
<i>BERTabaporu: Assessing a Genre-Specific Language Model for Portuguese NLP</i> Pablo Botton Costa, Matheus Camasmie Pavan, Wesley Ramos Santos, Samuel Caetano Silva and Ivandr� Paraboni .....	217
<i>Comparison of Multilingual Entity Linking Approaches</i> Ivelina Bozhinova and Andrey Tagarev .....	224
<i>Automatic Extraction of the Romanian Academic Word List: Data and Methods</i> Ana-Maria Bucur, Andreea Dinc�, Madalina Chitez and Roxana Rogobete .....	234
<i>Stance Prediction from Multimodal Social Media Data</i> Lais Carraro Leme Cavalheiro, Matheus Camasmie Pavan and Ivandr� Paraboni .....	242
<i>From Stigma to Support: A Parallel Monolingual Corpus and NLP Approach for Neutralizing Mental Illness Bias</i> Mason Choey .....	249
<i>BB25HLegalSum: Leveraging BM25 and BERT-Based Clustering for the Summarization of Legal Doc- uments</i> Leonardo de Andrade and Karin Becker .....	255
<i>SSSD: Leveraging Pre-trained Models and Semantic Search for Semi-supervised Stance Detection</i> Andr� de Sousa and Karin Becker .....	264
<i>Detecting Text Formality: A Study of Text Classification Approaches</i> Daryna Dementieva, Nikolay Babakov and Alexander Panchenko .....	274
<i>Developing a Multilingual Corpus of Wikipedia Biographies</i> Hannah Devinney, Anton Eklund, Igor Ryazanov and Jingwen Cai .....	285
<i>A Computational Analysis of the Voices of Shakespeare's Characters</i> Liviu P. Dinu and Ana Sabina Uban .....	295

<i>Source Code Plagiarism Detection with Pre-Trained Model Embeddings and Automated Machine Learning</i>	
Fahad Ebrahim and Mike Joy .....	301
<i>Identifying Semantic Argument Types in Predication and Copredication Contexts: A Zero-Shot Cross-Lingual Approach</i>	
Deniz Ekin Yavas, Laura Kallmeyer, Rainer Osswald, Elisabetta Jezeq, Marta Ricciardi and Long Chen .....	310
<i>A Review of Research-Based Automatic Text Simplification Tools</i>	
Isabel Espinosa-Zaragoza, José Abreu-Salas, Elena Lloret, Paloma Moreda and Manuel Palomar	321
<i>Vocab-Expander: A System for Creating Domain-Specific Vocabularies Based on Word Embeddings</i>	
Michael Faerber and Nicholas Popovic .....	331
<i>On the Generalization of Projection-Based Gender Debiasing in Word Embedding</i>	
Elisabetta Fersini, Antonio Candelieri and Lorenzo Pastore .....	336
<i>Mapping Explicit and Implicit Discourse Relations between the RST-DT and the PDTB 3.0</i>	
Nelson Filipe Costa, Nadia Sheikh and Leila Kosseim .....	344
<i>Bigfoot in Big Tech: Detecting Out of Domain Conspiracy Theories</i>	
Matthew Fort, Zuoyu Tian, Elizabeth Gabel, Nina Georgiades, Noah Sauer, Daniel Dakota and Sandra Kübler .....	353
<i>Deep Learning Approaches to Detecting Safeguarding Concerns in Schoolchildren’s Online Conversations</i>	
Emma Franklin and Tharindu Ranasinghe .....	364
<i>On the Identification and Forecasting of Hate Speech in Inceldom</i>	
Paolo Gajo, Arianna Muti, Katerina Korre, Silvia Bernardini and Alberto Barrón-Cedeño .....	373
<i>T2KG: Transforming Multimodal Document to Knowledge Graph</i>	
Santiago Galiano, Rafael Muñoz, Yoan Gutiérrez, Andrés Montoyo, Jose Ignacio Abreu and Luis Alfonso Ureña .....	385
<i>!Translate: When You Cannot Cook Up a Translation, Explain</i>	
Federico Garcea, Margherita Martinelli, Maja Milicević Petrović and Alberto Barrón-Cedeño .....	392
<i>An Evaluation of Source Factors in Concatenation-Based Context-Aware Neural Machine Translation</i>	
Harritxu Gete and Thierry Etchegoyhen .....	399
<i>Lessons Learnt from Linear Text Segmentation: a Fair Comparison of Architectural and Sentence Encoding Strategies for Successful Segmentation</i>	
Iacopo Ghinassi, Lin Wang, Chris Newell and Matthew Purver .....	408
<i>Student’s t-Distribution: On Measuring the Inter-Rater Reliability When the Observations are Scarce</i>	
Serge Gladkoff, Lifeng Han and Goran Nenadic .....	419
<i>Data Augmentation for Fake News Detection by Combining Seq2seq and NLI</i>	
Anna Glazkova .....	429

<i>Exploring Unsupervised Semantic Similarity Methods for Claim Verification in Health Care News Articles</i>	
Vishwani Gupta, Astrid Viciano, Holger Wormer and Najmehsadat Mousavinezhad . . . . .	440
<i>AlphaMWE-Arabic: Arabic Edition of Multilingual Parallel Corpora with Multiword Expression Annotations</i>	
najet hadj mohamed, Malak Rassem, Lifeng Han and Goran Nenadic . . . . .	448
<i>Performance Analysis of Arabic Pre-trained Models on Named Entity Recognition Task</i>	
Abdelhalim Hafedh Dahou, Mohamed Amine Cheragui and Ahmed Abdelali . . . . .	458
<i>Discourse Analysis of Argumentative Essays of English Learners Based on CEFR Level</i>	
Blaise Hanel and Leila Kosseim . . . . .	468
<i>Improving Translation Quality for Low-Resource Inuktitut with Various Preprocessing Techniques</i>	
Mathias Hans Erik Stenlund, Mathilde Nanni, Micaella Bruton and Meriem Beloucif . . . . .	475
<i>Enriched Pre-trained Transformers for Joint Slot Filling and Intent Detection</i>	
Momchil Hardalov, Ivan Koychev and Preslav Nakov . . . . .	480
<i>Unimodal Intermediate Training for Multimodal Meme Sentiment Classification</i>	
Muzhaffar Hazman, Susan McKeever and Josephine Griffith . . . . .	494
<i>Explainable Event Detection with Event Trigger Identification as Rationale Extraction</i>	
Hansi Hettiarachchi and Tharindu Ranasinghe . . . . .	507
<i>Clinical Text Classification to SNOMED CT Codes Using Transformers Trained on Linked Open Medical Ontologies</i>	
Anton Hristov, Petar Ivanov, Anna Aksenova, Tsvetan Asamov, Pavlin Gyurov, Todor Primov and Svetla Boytcheva . . . . .	519
<i>Towards a Consensus Taxonomy for Annotating Errors in Automatically Generated Text</i>	
Rudali Huidrom and Anya Belz . . . . .	527
<i>Uncertainty Quantification of Text Classification in a Multi-Label Setting for Risk-Sensitive Systems</i>	
Jinha Hwang, Carol Gudumotu and Benjamin Ahmadnia . . . . .	541
<i>Pretraining Language- and Domain-Specific BERT on Automatically Translated Text</i>	
Tatsuya Ishigaki, Yui Uehara, Goran Topić and Hiroya Takamura . . . . .	548
<i>Categorising Fine-to-Coarse Grained Misinformation: An Empirical Study of the COVID-19 Infodemic</i>	
Ye Jiang, Xingyi Song, Carolina Scarton, Iknor Singh, Ahmet Aker and Kalina Bontcheva . . . . .	556
<i>Bridging the Gap between Subword and Character Segmentation in Pretrained Language Models</i>	
Shun Kiyono, Sho Takase, Shengzhe Li and Toshinori Sato . . . . .	568
<i>Evaluating Data Augmentation for Medication Identification in Clinical Notes</i>	
Jordan Koontz, Maite Oronoz and Alicia Pérez . . . . .	578
<i>Advancing Topical Text Classification: A Novel Distance-Based Method with Contextual Embeddings</i>	
Andriy Kosar, Guy De Pauw and Walter Daelemans . . . . .	586
<i>Taxonomy-Based Automation of Prior Approval Using Clinical Guidelines</i>	
Saranya Krishnamoorthy and Ayush Singh . . . . .	598

<i>Simultaneous Interpreting as a Noisy Channel: How Much Information Gets Through</i> Maria Kunilovskaya, Heike Przybyl, Ekaterina Lapshinova-Koltunski and Elke Teich .....	608
<i>Challenges of GPT-3-Based Conversational Agents for Healthcare</i> Fabian Lechner, Allison Lahnala, Charles Welch and Lucie Flek .....	619
<i>Noisy Self-Training with Data Augmentations for Offensive and Hate Speech Detection Tasks</i> João Leite, Carolina Scarton and Diego Silva .....	631
<i>A Practical Survey on Zero-Shot Prompt Design for In-Context Learning</i> Yinheng Li .....	641
<i>Classifying COVID-19 Vaccine Narratives</i> Yue Li, Carolina Scarton, Xingyi Song and Kalina Bontcheva .....	648
<i>Sign Language Recognition and Translation: A Multi-Modal Approach Using Computer Vision and Natural Language Processing</i> Jacky Li, Jaren Gerdes, James Gojit, Austin Tao, Samyak Katke, Kate Nguyen and Benyamin Ahmadnia .....	658
<i>Classification-Aware Neural Topic Model Combined with Interpretable Analysis - for Conflict Classification</i> Tianyu Liang, Yida Mu, Soonho Kim, Darline Kuate, Julie Lang, Rob Vos and Xingyi Song ..	666
<i>Data Augmentation for Fake Reviews Detection</i> Ming Liu and Massimo Poesio .....	673
<i>Coherent Story Generation with Structured Knowledge</i> Congda Ma, Kotaro Funakoshi, Kiyooki Shirai and Manabu Okumura .....	681
<i>Studying Common Ground Instantiation Using Audio, Video and Brain Behaviours: The BrainKT Corpus</i> Eliot Maës, Thierry Legou, Leonor Becerra and Philippe Blache .....	691
<i>Reading between the Lines: Information Extraction from Industry Requirements</i> Ole Magnus Holter and Basil Ell .....	703
<i>Transformer-Based Language Models for Bulgarian</i> Iva Marinova, Kiril Simov and Petya Osenova .....	712
<i>Multi-task Ensemble Learning for Fake Reviews Detection and Helpfulness Prediction: A Novel Approach</i> Alimuddin Melleng, Anna Jurek-Loughrey and Deepak P. ....	721
<i>Data Fusion for Better Fake Reviews Detection</i> Alimuddin Melleng, Anna Jurek-Loughrey and Deepak P. ....	730
<i>Dimensions of Quality: Contrasting Stylistic vs. Semantic Features for Modelling Literary Quality in 9,000 Novels</i> Pascale Moreira and Yuri Bizzoni .....	739
<i>BanglaBait: Semi-Supervised Adversarial Approach for Clickbait Detection on Bangla Clickbait Dataset</i> Md. Motahar Mahtab, Monirul Haque, Mehedi Hasan and Farig Sadeque .....	748
<i>TreeSwap: Data Augmentation for Machine Translation via Dependency Subtree Swapping</i> Attila Nagy, Dorina Lakatos, Botond Barta and Judit Ács .....	759

<i>Automatic Assessment Of Spoken English Proficiency Based on Multimodal and Multitask Transformers</i> Kamel Nebhi and György Szaszák .....	769
<i>Medical Concept Mention Identification in Social Media Posts Using a Small Number of Sample References</i> Vasudevan Nedumpozhimana, Sneha Rautmare, Meegan Gower, Nishtha Jain, Maja Popović, Patricia Buffini and John Kelleher .....	777
<i>Context-Aware Module Selection in Modular Dialog Systems</i> Jan Nehring, René Marcel Berk and Stefan Hillmann .....	785
<i>Human Value Detection from Bilingual Sensory Product Reviews</i> Boyu Niu, Céline Manetta and Frédérique Segond .....	792
<i>Word Sense Disambiguation for Automatic Translation of Medical Dialogues into Pictographs</i> Magali Norré, Rémi Cardon, Vincent Vandeghinste and Thomas François .....	803
<i>A Research-Based Guide for the Creation and Deployment of a Low-Resource Machine Translation System</i> John E. Ortega and Kenneth Church .....	813
<i>MQDD: Pre-training of Multimodal Question Duplicity Detection for Software Engineering Domain</i> Jan Pasek, Jakub Sido, Miloslav Konopik and Ondrej Prazak .....	824
<i>Forming Trees with Treeformers</i> Nilay Patel and Jeffrey Flanigan .....	836
<i>Evaluating Unsupervised Hierarchical Topic Models Using a Labeled Dataset</i> Judicael Poumay and Ashwin Ittoo .....	846
<i>HTMOT: Hierarchical Topic Modelling over Time</i> Judicael Poumay and Ashwin Ittoo .....	854
<i>Multilingual Continual Learning Approaches for Text Classification</i> Karan Praharaj and Irina Matveeva .....	864
<i>Can Model Fusing Help Transformers in Long Document Classification? An Empirical Study</i> Damith Premasiri, Tharindu Ranasinghe and Ruslan Mitkov .....	871
<i>Deep Learning Methods for Identification of Multiword Flower and Plant Names</i> Damith Premasiri, Amal Haddad Haddad, Tharindu Ranasinghe and Ruslan Mitkov .....	879
<i>Improving Aspect-Based Sentiment with End-to-End Semantic Role Labeling Model</i> Pavel Přibáň and Ondrej Prazak .....	888
<i>huPWKP: A Hungarian Text Simplification Corpus</i> Noémi Prótár and Dávid Márk Nemeskey .....	898
<i>Topic Modeling Using Community Detection on a Word Association Graph</i> Mahfuzur Rahman Chowdhury, Intesur Ahmed, Farig Sadeque and Muhammad Yanhaona .....	908
<i>Exploring Techniques to Detect and Mitigate Non-Inclusive Language Bias in Marketing Communications Using a Dictionary-Based Approach</i> Bharathi Raja Chakravarthi, Prasanna Kumar Kumaresan, Rahul Ponnusamy, John P. McCrae, Michaela Comerford, Jay Megaro, Deniz Keles and Last Feremenga .....	918



<i>Does the "Most Sinfully Decadent Cake Ever" Taste Good? Answering Yes/No Questions from Figurative Contexts</i>	
Geetanjali Rakshit and Jeffrey Flanigan .....	926
<i>Modeling Easiness for Training Transformers with Curriculum Learning</i>	
Leonardo Ranaldi, Giulia Pucci and Fabio Massimo Zanzotto .....	937
<i>The Dark Side of the Language: Pre-trained Transformers in the DarkNet</i>	
Leonardo Ranaldi, Aria Nourbakhsh, Elena Sofia Ruzzetti, Arianna Patrizi, Dario Onorati, Michele Mastromattei, Francesca Fallucchi and Fabio Massimo Zanzotto .....	949
<i>PreCog: Exploring the Relation between Memorization and Performance in Pre-trained Language Models</i>	
Leonardo Ranaldi, Elena Sofia Ruzzetti and Fabio Massimo Zanzotto .....	961
<i>Publish or Hold? Automatic Comment Moderation in Luxembourgish News Articles</i>	
Tharindu Ranasinghe, Alistair Plum, Christoph Purschke and Marcos Zampieri .....	968
<i>Cross-Lingual Speaker Identification for Indian Languages</i>	
Amaan Rizvi, Anupam Jamatia, Dwijen Rudrapal, Kunal Chakma and Björn Gamb&auml;ck	979
<i>'ChemXtract' A System for Extraction of Chemical Events from Patent Documents</i>	
Pattabhi RK Rao and Sobha Lalitha Devi .....	988
<i>Mind the User! Measures to More Accurately Evaluate the Practical Value of Active Learning Strategies</i>	
Julia Romberg .....	996
<i>Event Annotation and Detection in Kannada-English Code-Mixed Social Media Data</i>	
Sumukh S, Abhinav Appidi and Manish Shrivastava .....	1007
<i>Three Approaches to Client Email Topic Classification</i>	
Branislava Šandrih Todorović, Katarina Josipović and Jurij Kodre .....	1015
<i>Exploring Abstractive Text Summarisation for Podcasts: A Comparative Study of BART and T5 Models</i>	
Parth Saxena and Mo El-Haj .....	1023
<i>Exploring the Landscape of Natural Language Processing Research</i>	
Tim Schopf, Karim Arabi and Florian Matthes .....	1034
<i>Efficient Domain Adaptation of Sentence Embeddings Using Adapters</i>	
Tim Schopf, Dennis N. Schneider and Florian Matthes .....	1046
<i>AspectCSE: Sentence Embeddings for Aspect-Based Semantic Textual Similarity Using Contrastive Learning and Structured Knowledge</i>	
Tim Schopf, Emanuel Gerber, Malte Ostendorff and Florian Matthes .....	1054
<i>Tackling the Myriads of Collusion Scams on YouTube Comments of Cryptocurrency Videos</i>	
Sadat Shahriar and Arjun Mukherjee .....	1066
<i>Exploring Deceptive Domain Transfer Strategies: Mitigating the Differences among Deceptive Domains</i>	
Sadat Shahriar, Arjun Mukherjee and Omprakash Gnawali .....	1076
<i>Party Extraction from Legal Contract Using Contextualized Span Representations of Parties</i>	
Sanjeevan Sivapiran, Charangan Vasantharajan and Uthayasanker Thayasivam .....	1085

<i>From Fake to Hyperpartisan News Detection Using Domain Adaptation</i> Răzvan-Alexandru Smădu, Sebastian-Vasile Echim, Dumitru-Clementin Cercel, Iuliana Marin and Florin Pop .....	1095
<i>Prompt-Based Approach for Czech Sentiment Analysis</i> Jakub Šmíd and Pavel Přibáň .....	1110
<i>Measuring Gender Bias in Natural Language Processing: Incorporating Gender-Neutral Linguistic Forms for Non-Binary Gender Identities in Abusive Speech Detection</i> Nasim Sobhani, Kinshuk Sengupta and Sarah Jane Delany .....	1121
<i>LeSS: A Computationally-Light Lexical Simplifier for Spanish</i> Sanja Stajner, Daniel Ibanez and Horacio Saggion .....	1132
<i>Hindi to Dravidian Language Neural Machine Translation Systems</i> Vijay Sundar Ram and Sobha Lalitha Devi .....	1143
<i>Looking for Traces of Textual Deepfakes in Bulgarian on Social Media</i> Irina Temnikova, Iva Marinova, Silvia Gargova, Ruslana Margova and Ivan Koychev .....	1151
<i>Propaganda Detection in Russian Telegram Posts in the Scope of the Russian Invasion of Ukraine</i> Natalia Vanetik, Marina Litvak, Egor Reviakin and Margarita Tiamanova .....	1162
<i>Auto-Encoding Questions with Retrieval Augmented Decoding for Unsupervised Passage Retrieval and Zero-Shot Question Generation</i> Stalin Varanasi, Muhammad Umer Tariq Butt and Guenter Neumann .....	1171
<i>NoHateBrazil: A Brazilian Portuguese Text Offensiveness Analysis System</i> Francielle Vargas, Isabelle Carvalho, Wolfgang Schmeisser-Nieto, Fabrício Benevenuto and Thiago Pardo .....	1180
<i>Socially Responsible Hate Speech Detection: Can Classifiers Reflect Social Stereotypes?</i> Francielle Vargas, Isabelle Carvalho, Ali Hürriyetoğlu, Thiago Pardo and Fabrício Benevenuto	1187
<i>Predicting Sentence-Level Factuality of News and Bias of Media Outlets</i> Francielle Vargas, Kokil Jaidka, Thiago Pardo and Fabrício Benevenuto .....	1197
<i>Classification of US Supreme Court Cases Using BERT-Based Techniques</i> Shubham Vatsal, Adam Meyers and John E. Ortega .....	1207
<i>Kāraka-Based Answer Retrieval for Question Answering in Indic Languages</i> Devika Verma, Ramprasad S. Joshi, Aiman A. Shivani and Rohan D. Gupta .....	1216
<i>Comparative Analysis of Named Entity Recognition in the Dungeons and Dragons Domain</i> gayashan Weerasundara and Nisansa de Silva .....	1225
<i>Comparative Analysis of Anomaly Detection Algorithms in Text Data</i> Yizhou Xu, Kata Gábor, Jérôme Milleret and Frédérique Segond .....	1234
<i>Poetry Generation Combining Poetry Theme Labels Representations</i> Yingyu Yan, Dongzhen Wen, Liang Yang, Dongyu Zhang and Hongfei LIN .....	1246
<i>Evaluating Generative Models for Graph-to-Text Generation</i> Shuzhou Yuan and Michael Faerber .....	1256

<i>Microsyntactic Unit Detection Using Word Embedding Models: Experiments on Slavic Languages</i>	
Iuliia Zaitova, Irina Stenger and Tania Avgustinova .....	1265
<i>Systematic TextRank Optimization in Extractive Summarization</i>	
Morris Zieve, Anthony Gregor, Frederik Juul Stokbaek, Hunter Lewis, Ellis Marie Mendoza and Benjamin Ahmadnia .....	1274