

2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)

Torino, Italy
20 – 25 May 2024

Volume 1 of 22

Editors:

**Nicoletta Calzolari
Min-Yen Kan
Veronique Hoste**

**Alessandro Lenci
Sakriani Sakti
Nianwen Xue**

ISBN: 978-1-7138-9760-6

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© (2024) by ELRA Language Resources Association
All rights reserved.

Copyright for individual papers remains with the authors and are licensed under a Creative Commons 4.0 license, CC-BY-ND. (<https://creativecommons.org/licenses/by-nd/4.0/>)

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

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>3AM: An Ambiguity-Aware Multi-Modal Machine Translation Dataset</i> Xinyu Ma, Xuebo Liu, Derek F. Wong, Jun Rao, Bei Li, Liang Ding, Lidia S. Chao, Dacheng Tao and Min Zhang	1
<i>A Benchmark Evaluation of Clinical Named Entity Recognition in French</i> Nesrine Bannour, Christophe Servan, Aurélie Névéol and Xavier Tannier	14
<i>A Benchmark for Recipe Understanding in Artificial Agents</i> Jens Nevens, Robin de Haes, Rachel Ringe, Mihai Pomarlan, Robert Porzel, Katrien Beuls and Paul van Eecke	22
<i>ABLE: Agency-BeLiefs Embedding to Address Stereotypical Bias through Awareness Instead of Obliviousness</i> Michelle YoungJin Kim, Junghwan Kim and Kristen Johnson	43
<i>Abstractive Multi-Video Captioning: Benchmark Dataset Construction and Extensive Evaluation</i> Rikito Takahashi, Hirokazu Kiyomaru, Chenhui Chu and Sadao Kurohashi	57
<i>Abstract-level Deductive Reasoning for Pre-trained Language Models</i> Xin Wu, Yi Cai and Ho-fung Leung	70
<i>A Call for Clarity in Beam Search: How It Works and When It Stops</i> Jungo Kasai, Keisuke Sakaguchi, Ronan Le Bras, Dragomir Radev, Yejin Choi and Noah A. Smith	77
<i>A Canonical Form for Flexible Multiword Expressions</i> Jan Odijk	91
<i>A Cause-Effect Look at Alleviating Hallucination of Knowledge-grounded Dialogue Generation</i> Jifan Yu, Xiaohan Zhang, Yifan Xu, Xuanyu Lei, Zijun Yao, Jing Zhang, Lei Hou and Juanzi Li	102
<i>Access Control Framework for Language Collections</i> Ben Foley, Peter Sefton, Simon Musgrave and Moises Sacal Bonequi	113
<i>A Challenge Dataset and Effective Models for Conversational Stance Detection</i> Fuqiang Niu, Min Yang, Ang Li, Baoquan Zhang, Xiaojiang Peng and Bowen Zhang ..	122
<i>A Closer Look at Clustering Bilingual Comparable Corpora</i> Anna Laskina, Eric Gaussier and Gaelle Calvary	133
<i>AcnEmpathize: A Dataset for Understanding Empathy in Dermatology Conversations</i> Gyeongeeun Lee and Natalie Parde	143
<i>A Collection of Pragmatic-Similarity Judgments over Spoken Dialog Utterances</i> Nigel Ward and Divette Marco	154
<i>A Community-Driven Data-to-Text Platform for Football Match Summaries</i> Pedro Fernandes, Sérgio Nunes and Luís Santos	164

<i>A Comparative Analysis of Word-Level Metric Differential Privacy: Benchmarking the Privacy-Utility Trade-off</i>	
Stephen Joseph Meisenbacher, Nihildev Nandakumar, Alexandra Klymenko and Florian Matthes	174
<i>A Comparative Study of Explicit and Implicit Gender Biases in Large Language Models via Self-evaluation</i>	
Yachao Zhao, Bo Wang, Yan Wang, Dongming Zhao, Xiaojia Jin, Jijun Zhang, Ruifang He and Yuexian Hou	186
<i>A Computational Analysis of the Dehumanisation of Migrants from Syria and Ukraine in Slovene News Media</i>	
Jaya Caporusso, Damar Hoogland, Mojca Brglez, Boshko Koloski, Matthew Purver and Senja Pollak	199
<i>A Computational Approach to Quantifying Grammaticization of English Deverbal Prepositions</i>	
Ryo Nagata, Yoshifumi Kawasaki, Naoki Otani and Hiroya Takamura	211
<i>A Computational Model of Latvian Morphology</i>	
Peteris Paikens, Lauma Pretkalniņa and Laura Rituma	221
<i>A Concept Based Approach for Translation of Medical Dialogues into Pictographs</i>	
Johanna Gerlach, Pierrette Bouillon, Jonathan Mutal and Hervé Spechbach	233
<i>A Construction Grammar Corpus of Varying Schematicity: A Dataset for the Evaluation of Abstractions in Language Models</i>	
Claire Bonial and Harish Tayyar Madabushi	243
<i>A Controlled Reevaluation of Coreference Resolution Models</i>	
Ian Porada, Xiyuan Zou and Jackie Chi Kit Cheung	256
<i>A Corpus and Method for Chinese Named Entity Recognition in Manufacturing</i>	
Ruiting Li, Peiyan Wang, Libang Wang, Danqingxin Yang and Dongfeng Cai	264
<i>A Corpus for Sentence-Level Subjectivity Detection on English News Articles</i>	
Francesco Antici, Federico Ruggeri, Andrea Galassi, Katerina Korre, Arianna Muti, Alessandra Bardi, Alice Fedotova and Alberto Barrón-Cedeño	273
<i>A Corpus of German Abstract Meaning Representation (DeAMR)</i>	
Christoph Otto, Jonas Groschwitz, Alexander Koller, Xiulin Yang and Lucia Donatelli ..	286
<i>A Corpus of Spontaneous L2 English Speech for Real-situation Speaking Assessment</i>	
Sylvain Coulange, Marie-Hélène Fries, Monica Masperi and Solange Rossato	293
<i>Action and Reaction Go Hand in Hand! a Multi-modal Dialogue Act Aided Sarcasm Identification</i>	
Mohit Singh Tomar, Tulika Saha, Abhisek Tiwari and Sriparna Saha	298
<i>Action-Concentrated Embedding Framework: This Is Your Captain Sign-tokening</i>	
Hyunwook Yu, Suhyeon Shin, Junku Heo, Hyuntaek Shin, Hyosu Kim and Mucheol Kim	310
<i>Active Learning Design Choices for NER with Transformers</i>	
Robert Vacareanu, Enrique Noriega-Atala, Gus Hahn-Powell, Marco A. Valenzuela-Escarcega and Mihai Surdeanu	321

<i>A CURATED CATALOG: Rethinking the Extraction of Pretraining Corpora for Mid-Resourced Languages</i>	
Jorge Palomar-Giner, Jose Javier Saiz, Ferran Espuña, Mario Mina, Severino Da Dalt, Joan Llop, Malte Ostendorff, Pedro Ortiz Suarez, Georg Rehm, Aitor Gonzalez-Agirre and Marta Villegas	335
<i>AdaKron: An Adapter-based Parameter Efficient Model Tuning with Kronecker Product</i>	
Marco Braga, Alessandro Raganato and Gabriella Pasi	350
<i>Adaptive Reinforcement Tuning Language Models as Hard Data Generators for Sentence Representation</i>	
Bo Xu, Yifei Wu, Shouang Wei, Ming Du and Hongya Wang	358
<i>Adaptive Simultaneous Sign Language Translation with Confident Translation Length Estimation</i>	
Tong Sun, Biao Fu, Cong Hu, Liang Zhang, Ruiquan Zhang, Xiaodong Shi, Jinsong Su and Yidong Chen	372
<i>A Dataset for Named Entity Recognition and Entity Linking in Chinese Historical Newspapers</i>	
Baptiste Blouin, Cécile Armand and Christian Henriot	385
<i>A Dataset for Pharmacovigilance in German, French, and Japanese: Annotating Adverse Drug Reactions across Languages</i>	
Lisa Raithel, Hui-Syuan Yeh, Shuntaro Yada, Cyril Grouin, Thomas Lavergne, Aurélie Névéol, Patrick Paroubek, Philippe Thomas, Tomohiro Nishiyama, Sebastian Möller, Eiji Aramaki, Yuji Matsumoto, Roland Roller and Pierre Zweigenbaum	395
<i>Adding SPICE to Life: Speaker Profiling in Multiparty Conversations</i>	
Shivani Kumar, Rishabh Gupta, Md. Shad Akhtar and Tanmoy Chakraborty	415
<i>ADEA: An Argumentative Dialogue Dataset on Ethical Issues Concerning Future A.I. Applications</i>	
Christian Hauptmann, Adrian Krenzer, Antonia Krause and Frank Puppe	426
<i>A Decade of Scholarly Research on Open Knowledge Graphs</i>	
Houcemeddine Turki, Abraham Toluwase Owodunni, Mohamed Ali Hadj Taieb, René Fabrice Bile and Mohamed Ben Aouicha	438
<i>A Differentiable Integer Linear Programming Solver for Explanation-Based Natural Language Inference</i>	
Mokanarangan Thayaparan, Marco Valentino and André Freitas	449
<i>A Document-Level Text Simplification Dataset for Japanese</i>	
Yoshinari Nagai, Teruaki Oka and Mamoru Komachi	459
<i>A Dual-View Approach to Classifying Radiology Reports by Co-Training</i>	
Yutong Han, Yan Yuan and Lili Mou	477
<i>Advancing Semi-Supervised Learning for Automatic Post-Editing: Data-Synthesis by Mask-Infilling with Erroneous Terms</i>	
Wonkee Lee, Seong-Hwan Heo and Jong-Hyeok Lee	484

<i>Advancing Topic Segmentation and Outline Generation in Chinese Texts: The Paragraph-level Topic Representation, Corpus, and Benchmark</i>	
Feng Jiang, Weihao Liu, Xiaomin Chu, Peifeng Li, Qiaoming Zhu and Haizhou Li	495
<i>A Family of Pretrained Transformer Language Models for Russian</i>	
Dmitry Zmitrovich, Aleksandr Abramov, Andrey Kalmykov, Vitaly Kadulin, Maria Tikhonova, Ekaterina Taktasheva, Danil Astafurov, Mark Baushenko, Artem Snegirev, Tatiana Shavrina, Sergei S. Markov, Vladislav Mikhailov and Alena Fenogenova	507
<i>A Fast and High-quality Text-to-Speech Method with Compressed Auxiliary Corpus and Limited Target Speaker Corpus</i>	
Ye Tao, Chaofeng Lu, Meng Liu, Kai Xu, Tianyu Liu, Yunlong Tian and Yongjie Du	525
<i>A Frustratingly Simple Decoding Method for Neural Text Generation</i>	
Haoran Yang, Deng Cai, Huayang Li, Wei Bi, Wai Lam and Shuming Shi	536
<i>A Gaze-grounded Visual Question Answering Dataset for Clarifying Ambiguous Japanese Questions</i>	
Shun Inadumi, Seiya Kawano, Akishige Yuguchi, Yasutomo Kawanishi and Koichiro Yoshino	558
<i>Agenda-Driven Question Generation: A Case Study in the Courtroom Domain</i>	
Yi Fung, Anoop Kumar, Aram Galstyan, Heng Ji and Prem Natarajan	572
<i>A Generative Model for Lambek Categorial Sequents</i>	
Jinman Zhao and Gerald Penn	584
<i>Agent-based Modeling of Language Change in a Small-world Network</i>	
Dalmo Buzato and Evandro Cunha	594
<i>Agettivu, Aggitivu o Aghjettivu? POS Tagging Corsican Dialects</i>	
Alice Millour, Lorenza Brasile, Alberto Ghia and Laurent Kevers	600
<i>Aggregation of Reasoning: A Hierarchical Framework for Enhancing Answer Selection in Large Language Models</i>	
Zhangyue Yin, Qiushi Sun, Qipeng Guo, Zhiyuan Zeng, Xiaonan Li, Tianxiang Sun, Cheng Chang, Qinyuan Cheng, Ding Wang, Xiaofeng Mou, Xipeng Qiu and Xuanjing Huang	609
<i>A Hierarchical Sequence-to-Set Model with Coverage Mechanism for Aspect Category Sentiment Analysis</i>	
Siyu Wang, Jianhui Jiang, Shengran Dai and Jiangtao Qiu	626
<i>A Hong Kong Sign Language Corpus Collected from Sign-interpreted TV News</i>	
Zhe Niu, Ronglai Zuo, Brian Mak and Fangyun Wei	636
<i>A Hybrid Approach to Aspect Based Sentiment Analysis Using Transfer Learning</i>	
Gaurav Negi, Rajdeep Sarkar, Omnia Zayed and Paul Buitelaar	647
<i>A Japanese News Simplification Corpus with Faithfulness</i>	
Toru Urakawa, Yuya Taguchi, Takuro Niitsuma and Hideaki Tamori	659
<i>A Knowledge Plug-and-Play Test Bed for Open-domain Dialogue Generation</i>	
Xiangci Li, Linfeng Song, Lifeng Jin, Haitao Mi, Jessica Ouyang and Dong Yu	666

<i>A Large Annotated Reference Corpus of New High German Poetry</i> Thomas Haider	677
<i>A Lifelong Multilingual Multi-granularity Semantic Alignment Approach via Maximum Co-occurrence Probability</i> Xin Liu, Hongwei Sun, Shaojie Dai, Bo Lv, Youcheng Pan, Hui Wang and Yue Yu.....	684
<i>A Lightweight Approach to a Giga-Corpus of Historical Periodicals: The Story of a Slovenian Historical Newspaper Collection</i> Filip Dobranić, Bojan Evkoski and Nikola Ljubešić	695
<i>Aligning the Norwegian UD Treebank with Entity and Coreference Information</i> Tollef Emil Jørgensen and Andre Kåsen	704
<i>Alignment before Awareness: Towards Visual Question Localized-Answering in Robotic Surgery via Optimal Transport and Answer Semantics</i> Zhihong Zhu, Yunyan Zhang, Xuxin Cheng, Zhiqi Huang, Derong Xu, Xian Wu and Yefeng Zheng	711
<i>Align-to-Distill: Trainable Attention Alignment for Knowledge Distillation in Neural Machine Translation</i> Heegon Jin, Seonil Son, Jemin Park, Youngseok Kim, Hyungjong Noh and Yeonsoo Lee	722
<i>A Linguistically-Informed Annotation Strategy for Korean Semantic Role Labeling</i> Yige Chen, KyungTae Lim and Jungyeul Park	733
<i>Alleviating Exposure Bias in Abstractive Summarization via Sequentially Generating and Revising</i> Jiaxin Duan, Fengyu Lu and Junfei Liu	739
<i>ALLIES: A Speech Corpus for Segmentation, Speaker Diarization, Speech Recognition and Speaker Change Detection</i> Marie Tahon, Anthony Larcher, Martin Lebourdais, Fethi Bougares, Anna Silnova and Pablo Gimeno	751
<i>A Logical Pattern Memory Pre-trained Model for Entailment Tree Generation</i> Li Yuan, Yi Cai, Haopeng Ren and Jiexin Wang	759
<i>AlphaFin: Benchmarking Financial Analysis with Retrieval-Augmented Stock-Chain Framework</i> Xiang Li, Zhenyu Li, Chen Shi, Yong Xu, Qing Du, Mingkui Tan and Jun Huang	773
<i>A Luxembourgish Corpus as a Gender Bias Evaluation Testset</i> Dimitra Anastasiou, Carole Blond-Hanten and Marie Gallais	784
<i>A Matter of Perspective: Building a Multi-Perspective Annotated Dataset for the Study of Literary Quality</i> Yuri Bizzoni, Pascale Feldkamp Moreira, Ida Marie S. Lassen, Mads Rosendahl Thomsen and Kristoffer Nielbo	789
<i>AMenDeD: Modelling Concepts by Aligning Mentions, Definitions and Decontextualised Embeddings</i> Amit Gajbhiye, Zied Bouraoui, Luis Espinosa Anke and Steven Schockaert	801

<i>A Multi-Label Dataset of French Fake News: Human and Machine Insights</i> Benjamin Icard, François Maine, Morgane Casanova, Géraud Faye, Julien Chanson, Guillaume Gadek, Ghislain Ateazing, François Bancilhon and Paul Égré	812
<i>A Multi-layered Approach to Physical Commonsense Understanding: Creation and Evaluation of an Italian Dataset</i> Giulia Pensa, Begoña Altuna and Itziar Gonzalez-Dios	819
<i>A Multilingual Parallel Corpus for Aromanian</i> Iulia Petrariu and Sergiu Nisioi	832
<i>A Multimodal French Corpus of Aligned Speech, Text, and Pictogram Sequences for Speech-to-Pictogram Machine Translation</i> Cécile Macaire, Chloé Dion, Jordan Arrigo, Claire Lemaire, Emmanuelle Esperança-Rodier, Benjamin Lecouteux and Didier Schwab	839
<i>A Multimodal In-Context Tuning Approach for E-Commerce Product Description Generation</i> Yunxin Li, Baotian Hu, Wenhan Luo, Lin Ma, Yuxin Ding and Min Zhang	850
<i>A Multi-Task Transformer Model for Fine-grained Labelling of Chest X-Ray Reports</i> Yuanyi Zhu, Maria Liakata and Giovanni Montana	862
<i>Analysis of Sensation-transfer Dialogues in Motorsports</i> Takeru Isaka, Atsushi Otsuka and Iwaki Toshima	876
<i>Analysis on Unsupervised Acquisition Process of Bilingual Vocabulary through Iterative Back-Translation</i> Takuma Tanigawa, Tomoyosi Akiba and Hajime Tsukada	887
<i>Analyzing Chain-of-thought Prompting in Black-Box Large Language Models via Estimated V-information</i> Zecheng Wang, Chunshan Li, Zhao Yang, Qingbin Liu, Yanchao Hao, Xi Chen, Dianhui Chu and Dianbo Sui	893
<i>Analyzing Effects of Learning Downstream Tasks on Moral Bias in Large Language Models</i> Niklas Kiehne, Alexander Ljapunov, Marc Bätje and Wolf-Tilo Balke	904
<i>Analyzing Homonymy Disambiguation Capabilities of Pretrained Language Models</i> Lorenzo Proietti, Stefano Perrella, Simone Tedeschi, Giulia Vulpis, Leonardo Lavalle, Andrea Sanchietti, Andrea Ferrari and Roberto Navigli	924
<i>Analyzing Interpretability of Summarization Model with Eye-gaze Information</i> Fariz Ikhwantri, Hiroaki Yamada and Takenobu Tokunaga	939
<i>Analyzing Large Language Models' Capability in Location Prediction</i> Zhaomin Xiao, Eduardo Blanco and Yan Huang	951
<i>Analyzing Occupational Distribution Representation in Japanese Language Models</i> Katsumi Ibaraki, Winston Wu, Lu Wang and Rada Mihalcea	959
<i>Analyzing Symptom-based Depression Level Estimation through the Prism of Psychiatric Expertise</i> Navneet Agarwal, Kirill Milintsevich, Lucie Metivier, Maud Rotharmel, Gaël Dias and Sonia Dollfus	974

<i>Analyzing the Dynamics of Climate Change Discourse on Twitter: A New Annotated Corpus and Multi-Aspect Classification</i>	
Shuvam Shiwakoti, Surendrabikram Thapa, Kritesh Rauniyar, Akshyat Shah, Aashish Bhandari and Usman Naseem	984
<i>Analyzing the Performance of Large Language Models on Code Summarization</i>	
Rajarshi Haldar and Julia Hockenmaier	995
<i>Analyzing the Understanding of Morphologically Complex Words in Large Language Models</i>	
Marion Weller-Di Marco and Alexander Fraser	1009
<i>An Argument for Symmetric Coordination from Dependency Length Minimization: A Replication Study</i>	
Adam Przepiórkowski, Magdalena Borysiak and Adam Głowacki	1021
<i>A Natural Approach for Synthetic Short-Form Text Analysis</i>	
Ruiting Shao, Ryan Schwarz, Christopher Clifton and Edward Delp	1034
<i>An Automated End-to-End Open-Source Software for High-Quality Text-to-Speech Dataset Generation</i>	
Ahmet Gunduz, Kamer Ali Yuksel, Kareem Darwish, Golara Javadi, Fabio Minazzi, Nicola Sobieski and Sébastien Bratières	1043
<i>Anchor and Broadcast: An Efficient Concept Alignment Approach for Evaluation of Semantic Graphs</i>	
Haibo Sun and Nianwen Xue	1052
<i>An Effective Span-based Multimodal Named Entity Recognition with Consistent Cross-Modal Alignment</i>	
Yongxiu Xu, Hao Xu, Heyan Huang, Shiyao Cui, Minghao Tang, Longzheng Wang and Hongbo Xu	1063
<i>An Empirical Study of Synthetic Data Generation for Implicit Discourse Relation Recognition</i>	
Kazumasa Omura, Fei Cheng and Sadao Kurohashi	1073
<i>An Empirical Study on the Robustness of Massively Multilingual Neural Machine Translation</i>	
Supryadi Supryadi, Leiyu Pan and Deyi Xiong	1086
<i>An Evaluation of Croatian ASR Models for Čakavian Transcription</i>	
Shulin Zhang, John Hale, Margaret Renwick, Zvezdana Vrzić and Keith Langston ..	1098
<i>An Event-based Abductive Learning for Hard Time-sensitive Question Answering</i>	
Shaojuan Wu, Jitong Li, Xiaowang Zhang and Zhiyong Feng	1105
<i>A New Massive Multilingual Dataset for High-Performance Language Technologies</i>	
Ona de Gibert, Graeme Nail, Nikolay Arefyev, Marta Bañón, Jelmer van der Linde, Shaoxiong Ji, Jaume Zaragoza-Bernabeu, Mikko Aulamo, Gema Ramírez-Sánchez, Andrey Kutuzov, Sampo Pyysalo, Stephan Oepen and Jörg Tiedemann	1116
<i>An LCF-IDF Document Representation Model Applied to Long Document Classification</i>	
Renzo Arturo Alva Principe, Nicola Chiarini and Marco Viviani	1129
<i>An LLM-Enhanced Adversarial Editing System for Lexical Simplification</i>	
Keren Tan, Kangyang Luo, Yunshi Lan, Zheng Yuan and Jinlong Shu	1136

<i>AnnoCTR: A Dataset for Detecting and Linking Entities, Tactics, and Techniques in Cyber Threat Reports</i>	
Lukas Lange, Marc Müller, Ghazaleh Haratinezhad Torbati, Dragan Milchevski, Patrick Grau, Subhash Chandra Pujari and Annemarie Friedrich	1147
<i>Annotate Chinese Aspect with UMR—a Case Study on the Little Prince</i>	
Sijia Ge, Zilong Li, Alvin Po-Chun Chen and Guanchao Wang.....	1161
<i>Annotate the Way You Think: An Incremental Note Generation Framework for the Summarization of Medical Conversations</i>	
Longxiang Zhang, Caleb D. Hart, Susanne Burger and Thomas Schaaf.....	1173
<i>Annotating Chinese Word Senses with English WordNet: A Practice on OntoNotes Chinese Sense Inventories</i>	
Hongzhi Xu, Jingxia Lin, Sameer Pradhan, Mitchell Marcus and Ming Liu	1187
<i>Annotating Customer-Oriented Behaviour in Call Centre Sales Dialogues</i>	
Jutta Stock, Volha Petukhova and Dietrich Klakow	1197
<i>Annotation and Classification of Relevant Clauses in Terms-and-Conditions Contracts</i>	
Pietro Giovanni Bizzaro, Elena Della Valentina, Maurizio Napolitano, Nadia Mana and Massimo Zancanaro.....	1209
<i>Annotation of Japanese Discourse Relations Focusing on Concessive Inferences</i>	
Ai Kubota, Takuma Sato, Takayuki Amamoto, Ryota Akiyoshi and Koji Mineshima ...	1215
<i>Annotation of Transition-Relevance Places and Interruptions for the Description of Turn-Taking in Conversations in French Media Content</i>	
Rémi Uro, Marie Tahon, Jane Wottawa, David Doukhan, Albert Rilliard and Antoine Laurent	1225
<i>Annotations for Exploring Food Tweets from Multiple Aspects</i>	
Matiss Rikters, Rinalds Vīksna and Edison Marrese-Taylor	1233
<i>Annotations on a Budget: Leveraging Geo-Data Similarity to Balance Model Performance and Annotation Cost</i>	
Oana Ignat, Longju Bai, Joan C. Nwatu and Rada Mihalcea.....	1239
<i>AnnoTheia: A Semi-Automatic Annotation Toolkit for Audio-Visual Speech Technologies</i>	
José-M. Acosta-Triana, David Gimeno-Gómez and Carlos-D. Martínez-Hinarejos....	1260
<i>Announcing the Prague Discourse Treebank 3.0</i>	
Pavλίna Synková, Jiří Mírovský, Lucie Poláková and Magdaléna Rysová.....	1270
<i>A Novel Corpus of Annotated Medical Imaging Reports and Information Extraction Results Using BERT-based Language Models</i>	
Namu Park, Kevin Lybarger, Giridhar Kaushik Ramachandran, Spencer Lewis, Aashka Damani, Özlem Uzuner, Martin Gunn and Meliha Yetisgen.....	1280
<i>A Novel Three-stage Framework for Few-shot Named Entity Recognition</i>	
Shengjie Ji and Fang Kong	1293
<i>AntCritic: Argument Mining for Free-Form and Visually-Rich Financial Comments</i>	
Huadai Liu, Xu Wenqiang, Xuan Lin, Jingjing Huo, Hong Chen and Zhou Zhao	1306

<i>An Unsupervised Framework for Adaptive Context-aware Simplified-Traditional Chinese Character Conversion</i>	
Wei Li, Shutan Huang and Yanqiu Shao	1318
<i>An Untold Story of Preprocessing Task Evaluation: An Alignment-based Joint Evaluation Approach</i>	
Eunkyul Leah Jo, Angela Yoonseo Park, Grace Tianjiao Zhang, Izia Xiaoxiao Wang, Junrui Wang, MingJia Mao and Jungyeul Park.....	1327
<i>A Paradigm Shift: The Future of Machine Translation Lies with Large Language Models</i>	
Chenyang Lyu, Zefeng Du, Jitao Xu, Yitao Duan, Minghao Wu, Teresa Lynn, Alham Fikri Aji, Derek F. Wong and Longyue Wang.....	1339
<i>A Persona-Based Corpus in the Diabetes Self-Care Domain - Applying a Human-Centered Approach to a Low-Resource Context</i>	
Rossana Cunha, Thiago Castro Ferreira, Adriana Pagano and Fabio Alves.....	1353
<i>APOLLO: An Optimized Training Approach for Long-form Numerical Reasoning</i>	
Jiashuo Sun, Hang Zhang, Chen Lin, Xiangdong Su, Yeyun Gong and Jian Guo	1370
<i>Applying Transfer Learning to German Metaphor Prediction</i>	
Maria Berger, Sebastian Michael Reimann and Nieke Marie Kiwitt	1383
<i>Appraisal Framework for Clinical Empathy: A Novel Application to Breaking Bad News Conversations</i>	
Allison Claire Lahnala, Béla Neuendorf, Alexander Thomin, Charles Welch, Tina Stibane and Lucie Flek.....	1393
<i>Approaches and Challenges for Resolving Different Representations of Fictional Characters for Chinese Novels</i>	
Li Song and Ying Liu	1408
<i>A Preliminary Study of ChatGPT for Spanish E2R Text Adaptation</i>	
Margot Madina, Itziar Gonzalez-Dios and Melanie Siegel	1422
<i>A Quantum-Inspired Matching Network with Linguistic Theories for Metaphor Detection</i>	
Wenbo Qiao, Peng Zhang and ZengLai Ma	1435
<i>Arabic Diacritization Using Morphologically Informed Character-Level Model</i>	
Muhammad Morsy Elmallah, Mahmoud Reda, Kareem Darwish, Abdelrahman El-Sheikh, Ashraf Hatim Elneima, Murtadha Aljubran, Nouf Alsaeed, Reem Mohammed and Mohamed Al-Badrashiny	1446
<i>Arbitrary Time Information Modeling via Polynomial Approximation for Temporal Knowledge Graph Embedding</i>	
Zhiyu Fang, Jingyan Qin, Xiaobin Zhu, Chun Yang and Xu-Cheng Yin.....	1455
<i>ARBRES Kenstur: A Breton-French Parallel Corpus Rooted in Field Linguistics</i>	
Loïc Grobol and Mélanie Jouitteau.....	1466
<i>A Regularization-based Transfer Learning Method for Information Extraction via Instructed Graph Decoder</i>	
Kedi Chen, Jie Zhou, Qin Chen, Shunyu Liu and Liang He	1472

<i>A Reinforcement Learning Approach to Improve Low-Resource Machine Translation Leveraging Domain Monolingual Data</i>	
Hongxiao Zhang, Mingtong Liu, Chunyou Li, Yufeng Chen, Jinan Xu and Ming Zhou	1486
<i>Are Large Language Models Good at Lexical Semantics? A Case of Taxonomy Learning</i>	
Viktor Moskvoretskii, Alexander Panchenko and Irina Nikishina	1498
<i>Are Text Classifiers Xenophobic? A Country-Oriented Bias Detection Method with Least Confounding Variables</i>	
Valentin Barriere and Sebastian Cifuentes	1511
<i>Argument Quality Assessment in the Age of Instruction-Following Large Language Models</i>	
Henning Wachsmuth, Gabriella Lapesa, Elena Cabrio, Anne Lauscher, Joonsuk Park, Eva Maria Vecchi, Serena Villata and Timon Ziegenbein	1519
<i>Article Classification with Graph Neural Networks and Multigraphs</i>	
Khang Ly, Yury Kashnitsky, Savvas Chamezopoulos and Valeria Krzhizhanovskaya	1539
<i>ART: The Alternating Reading Task Corpus for Speech Entrainment and Imitation</i>	
Zheng Byron Yuan, Dorina de Jong, Ruitao Feng, Štefan Beňuš, Noël Nguyen, Róbert Sabo, Luciano Fadiga and Alessandro D'Ausilio	1548
<i>A Self-verified Method for Exploring Simile Knowledge from Pre-trained Language Models</i>	
Longxuan Ma, Changxin Ke, Shuhan Zhou, Churui Sun, Wei-Nan Zhang and Ting Liu	1563
<i>A Semantic Mention Graph Augmented Model for Document-Level Event Argument Extraction</i>	
Jian Zhang, Changlin Yang, Haiping Zhu, Qika Lin, Fangzhi Xu and Jun Liu	1577
<i>ASEM: Enhancing Empathy in Chatbot through Attention-based Sentiment and Emotion Modeling</i>	
Omama Hamad, Khaled Shaban and Ali Hamdi	1588
<i>A Single Linear Layer Yields Task-Adapted Low-Rank Matrices</i>	
Hwichan Kim, Shota Sasaki, Sho Hoshino and Ukyo Honda	1602
<i>Asking and Answering Questions to Extract Event-Argument Structures</i>	
Md Nayem Uddin, Enfa Rose George, Eduardo Blanco and Steven R. Corman	1609
<i>AssameseBackTranslit: Back Transliteration of Romanized Assamese Social Media Text</i>	
Hemanta Baruah, Sanasam Ranbir Singh and Priyankoo Sarmah	1627
<i>Assessing Online Writing Feedback Resources: Generative AI vs. Good Samaritans</i>	
Shabnam Behzad, Omid Kashefi and Swapna Somasundaran	1638
<i>Assessing the Capabilities of Large Language Models in Coreference: An Evaluation</i>	
Yujian Gan, Massimo Poesio and Juntao Yu	1645
<i>Assessing the Efficacy of Grammar Error Correction: A Human Evaluation Approach in the Japanese Context</i>	
Qiao Wang and Zheng Yuan	1666
<i>A Streamlined Span-based Factorization Method for Few Shot Named Entity Recognition</i>	
Wenjie Xu, Yidan Chen and Jianquan Ouyang	1673

<i>A Study on How Attention Scores in the BERT Model Are Aware of Lexical Categories in Syntactic and Semantic Tasks on the GLUE Benchmark</i>	
Dongjun Jang, Sungjoo Byun and Hyopil Shin	1684
<i>A Survey on Natural Language Processing for Programming</i>	
Qingfu Zhu, Xianzhen Luo, Fang Liu, Cuiyun Gao and Wanxiang Che	1690
<i>A Tool for Determining Distances and Overlaps between Multimodal Annotations</i>	
Camila Antonio Barros, Jorge Francisco Ciprián-Sánchez and Saulo Mendes Santos	1705
<i>A Treebank of Asia Minor Greek</i>	
Eleni Vligouridou, Inessa Iliadou and Çağrı Çöltekin	1715
<i>A Trusted Multi-View Evidential Fusion Framework for Commonsense Reasoning</i>	
Shuo Yang	1722
<i>Attack Named Entity Recognition by Entity Boundary Interference</i>	
Yifei Yang, Hongqiu Wu and Hai Zhao	1734
<i>At the Crossroad of Cuneiform and NLP: Challenges for Fine-grained Part-of-speech Tagging</i>	
Gustav Ryberg Smidt, Els Lefever and Katrien de Graef	1745
<i>A Tulu Resource for Machine Translation</i>	
Manu Narayanan and Noëmi Aepli	1756
<i>A Two-Stage Framework with Self-Supervised Distillation for Cross-Domain Text Classification</i>	
Yunlong Feng, Bohan Li, Libo Qin, Xiao Xu and Wanxiang Che	1768
<i>A Two-Stage Prediction-Aware Contrastive Learning Framework for Multi-Intent NLU</i>	
Guanhua Chen, Yutong Yao, Derek F. Wong and Lidia S. Chao	1778
<i>A Typology of Errors for User Utterances in Chatbots</i>	
Anu Singh and Esme Manandise	1789
<i>Audiocite.net : A Large Spoken Read Dataset in French</i>	
Soline Felice, Solene Virginie Evain, Solange Rossato and François Portet	1795
<i>AuRoRA: A One-for-all Platform for Augmented Reasoning and Refining with Task-Adaptive Chain-of-Thought Prompting</i>	
Anni Zou, Zhuosheng Zhang and Hai Zhao	1801
<i>Automated Extraction of Prosodic Structure from Unannotated Sign Language Video</i>	
Antonio F. G. Sevilla, José María Lahoz-Bengoechea and Alberto Diaz	1808
<i>Automatically Estimating Textual and Phonemic Complexity for Cued Speech: How to See the Sounds from French Texts</i>	
Núria Gala, Brigitte Bigi and Marie Bauer	1817
<i>Automatic Animacy Classification for Romanian Nouns</i>	
Maria Tepei and Jelke Bloem	1825
<i>Automatic Annotation of Grammaticality in Child-Caregiver Conversations</i>	
Mitja Nikolaus, Abhishek Agrawal, Petros Kaklamanis, Alex Warstadt and Abdellah Fourtassi	1832

<i>Automatic Authorship Analysis in Human-AI Collaborative Writing</i> Aquia Richburg, Calvin Bao and Marine Carpuat	1845
<i>Automatic Coding of Contingency in Child-Caregiver Conversations</i> Abhishek Agrawal, Mitja Nikolaus, Benoit Favre and Abdellah Fourtassi	1856
<i>Automatic Construction of a Chinese Review Dataset for Aspect Sentiment Triplet Extraction via Iterative Weak Supervision</i> Chia-Wen Lu, Ching-Wen Yang and Wei-Yun Ma	1871
<i>Automatic Construction of a Large-Scale Corpus for Geoparsing Using Wikipedia Hyperlinks</i> Keyaki Ohno, Hirotaka Kameko, Keisuke Shirai, Taichi Nishimura and Shinsuke Mori	1883
<i>Automatic Data Visualization Generation from Chinese Natural Language Questions</i> Yan Ge, Victor Junqiu Wei, Yuanfeng Song, Jason Chen Zhang and Raymond Chi-Wing Wong	1889
<i>Automatic Decomposition of Text Editing Examples into Primitive Edit Operations: Toward Analytic Evaluation of Editing Systems</i> Daichi Yamaguchi, Rei Miyata, Atsushi Fujita, Tomoyuki Kajiwara and Satoshi Sato .	1899
<i>Automatic Extraction of Language-Specific Biomarkers of Healthy Aging in Icelandic</i> Elena Callegari, Iris Edda Nowenstein, Ingunn Jóhanna Kristjánsdóttir and Anton Karl Ingason	1915
<i>Automatic Extraction of Nominal Phrases from German Learner Texts of Different Proficiency Levels</i> Ronja Laarmann-Quante, Marco Müller and Eva Belke	1925
<i>Automatic Identification of COVID-19-Related Conspiracy Narratives in German Telegram Channels and Chats</i> Philipp Heinrich, Andreas Blombach, Bao Minh Doan Dang, Leonardo Zilio, Linda Havenstein, Nathan Dykes, Stephanie Evert and Fabian Schäfer	1932
<i>Automatic Partitioning of a Code-Switched Speech Corpus Using Mixed-Integer Programming</i> Joshua Miles Jansen van Vüren, Febe de Wet and Thomas Niesler	1944
<i>Automatic Punctuation Model for Spanish Live Transcriptions</i> Mario Perez-Enriquez, Jose Manuel Masiello-Ruiz, Jose Luis Lopez-Cuadrado, Israel Gonzalez-Carrasco, Paloma Martinez-Fernandez and Belen Ruiz-Mezcua	1953
<i>Automatic Speech Interruption Detection: Analysis, Corpus, and System</i> Martin Lebourdais, Marie Tahon, Antoine Laurent and Sylvain Meignier	1959
<i>Automatic Speech Recognition for Gascon and Languedocian Variants of Occitan</i> Iñigo Morcillo, Igor Leturia, Ander Corral, Xabier Sarasola, Michaël Barret, Aure Séguier and Benaset Dazéas	1969
<i>Automatic Speech Recognition System-Independent Word Error Rate Estimation</i> Chanho Park, Mingjie Chen and Thomas Hain	1979
<i>Automating Dataset Production Using Generative Text and Image Models</i> Christopher Thierauf, Mitchell Abrams and Matthias Scheutz	1988

<i>Autonomous Aspect-Image Instruction a2II: Q-Former Guided Multimodal Sentiment Classification</i>	1996
Junjia Feng, Mingqian Lin, Lin Shang and Xiaoying Gao	
<i>Auxiliary Knowledge-Induced Learning for Automatic Multi-Label Medical Document Classification</i>	2006
Xindi Wang, Robert E. Mercer and Frank Rudzicz	
<i>A Virtual Patient Dialogue System Based on Question-Answering on Clinical Records</i>	2017
Janire Arana, Mikel Idoyaga, Maitane Urruela, Elisa Espina, Aitziber Atutxa Salazar and Koldo Gojenola	
<i>A Web Portal about the State of the Art of NLP Tasks in Spanish</i>	2028
Enrique Amigó, Jorge Carrillo-de-Albornoz, Andrés Fernández, Julio Gonzalo, Guillermo Marco, Roser Morante, Laura Plaza and Jacobo Pedrosa	
<i>A Workflow for HTR-Postprocessing, Labeling and Classifying Diachronic and Regional Variation in Pre-Modern Slavic Texts</i>	2039
Piroska Lendvai, Maarten van Gompel, Anna Jouravel, Elena Renje, Uwe Reichel, Achim Rabus and Eckhart Arnold	
<i>A Zero-shot and Few-shot Study of Instruction-Finetuned Large Language Models Applied to Clinical and Biomedical Tasks</i>	2049
Yanis Labrak, Mickael Rouvier and Richard Dufour	
<i>Backdoor NLP Models via AI-Generated Text</i>	2067
Wei Du, Tianjie Ju, Ge Ren, GaoLei Li and Gongshen Liu	
<i>BalsuTalka.lv - Boosting the Common Voice Corpus for Low-Resource Languages</i>	2080
Roberts Dargis, Arturs Znotins, Ilze Auzina, Baiba Saulite, Sanita Reinsone, Raivis Dejus, Antra Klavinska and Normunds Gruzitis	
<i>BAMBOO: A Comprehensive Benchmark for Evaluating Long Text Modeling Capacities of Large Language Models</i>	2086
Zican Dong, Tianyi Tang, Junyi Li, Wayne Xin Zhao and Ji-Rong Wen	
<i>BanglaAutoKG: Automatic Bangla Knowledge Graph Construction with Semantic Neural Graph Filtering</i>	2100
Azmine Toushik Wasi, Taki Hasan Rafi, Raima Islam and Dong-Kyu Chae	
<i>BAN-PL: A Polish Dataset of Banned Harmful and Offensive Content from Wykop.pl Web Service</i>	2107
Anna Kolos, Inez Okulska, Kinga Głąbińska, Agnieszka Karlinska, Emilia Wisnios, Paweł Ellerik and Andrzej Prałat	
<i>"Barking up the Right Tree", a GAN-Based Pun Generation Model through Semantic Pruning</i>	2119
JingJie Zeng, Liang Yang, Jiahao Kang, Yufeng Diao, Zhihao Yang and Hongfei Lin	
<i>Basque and Spanish Counter Narrative Generation: Data Creation and Evaluation</i>	2132
Jaione Bengoetxea, Yi-Ling Chung, Marco Guerini and Rodrigo Agerri	
<i>Becoming a High-Resource Language in Speech: The Catalan Case in the Common Voice Corpus</i>	2142
Carme Armentano-Oller, Montserrat Marimon and Marta Villegas	

<i>BEIR-PL: Zero Shot Information Retrieval Benchmark for the Polish Language</i> Konrad Wojtasik, Kacper Wołowiec, Vadim Shishkin, Arkadiusz Janz and Maciej Piasecki	2149
<i>Benchmarking GPT-4 on Algorithmic Problems: A Systematic Evaluation of Prompting Strategies</i> Flavio Petruzzellis, Alberto Testolin and Alessandro Sperduti	2161
<i>Benchmarking Hallucination in Large Language Models Based on Unanswerable Math Word Problem</i> YuHong Sun, Zhangyue Yin, Qipeng Guo, Jiawen Wu, Xipeng Qiu and Hui Zhao	2178
<i>Benchmarking Large Language Models for Persian: A Preliminary Study Focusing on ChatGPT</i> Amirhossein Abaskohi, Sara Baruni, Mostafa Masoudi, Nesa Abbasi, Mohammad Hadi Babalou, Ali Edalat, Sepehr Kamahi, Samin Mahdizadeh Sani, nikoo naghavian, Danial Namazifard, Pouya Sadeghi and Yadollah Yaghoobzadeh	2189
<i>Benchmarking the Performance of Machine Translation Evaluation Metrics with Chinese Multi-word Expressions</i> Huacheng Song and Hongzhi Xu	2204
<i>Benchmarking the Simplification of Dutch Municipal Text</i> Daniel Vlantis, Iva Gornishka and Shuai Wang	2217
<i>BengaliLCP: A Dataset for Lexical Complexity Prediction in the Bengali Texts</i> Nabila Ayman, Md. Akram Hossain, Abdul Aziz, Rokan Uddin Faruqui and Abu Nowshed Chy	2227
<i>BenLLM-Eval: A Comprehensive Evaluation into the Potentials and Pitfalls of Large Language Models on Bengali NLP</i> Mohsinul Kabir, Mohammed Saidul Islam, Md Tahmid Rahman Laskar, Mir Tafseer Nayeem, M Saiful Bari and Enamul Hoque	2238
<i>BERT-BC: A Unified Alignment and Interaction Model over Hierarchical BERT for Response Selection</i> Zhenfei Yang, Beiming Yu, Yuan Cui, Shi Feng, Daling Wang and Yifei Zhang	2253
<i>Beyond Binary: Towards Embracing Complexities in Cyberbullying Detection and Intervention - a Position Paper</i> Kanishk Verma, Kolawole John Adebayo, Joachim Wagner, Megan Reynolds, Rebecca Umbach, Tijana Milosevic and Brian Davis	2264
<i>Beyond Canonical Fine-tuning: Leveraging Hybrid Multi-Layer Pooled Representations of BERT for Automated Essay Scoring</i> Eujene Nikka V. Boquio and Prospero C. Naval, Jr.	2285
<i>Beyond Code: Evaluate Thought Steps for Complex Code Generation</i> Liuwen Cao, Yi Cai, Jiexin Wang, Hongkui He and Hailin Huang	2296
<i>Beyond Full Fine-tuning: Harnessing the Power of LoRA for Multi-Task Instruction Tuning</i> Chunlei Xin, Yaojie Lu, Hongyu Lin, Shuheng Zhou, Huijia Zhu, Weiqiang Wang, Zhongyi Liu, Xianpei Han and Le Sun	2307

<i>Beyond Linguistic Cues: Fine-grained Conversational Emotion Recognition via Belief-Desire Modelling</i>	
Bo Xu, Longjiao Li, Wei Luo, Mehdi Naseriparsa, Zhehuan Zhao, Hongfei Lin and Feng Xia	2318
<i>Beyond Model Performance: Can Link Prediction Enrich French Lexical Graphs?</i>	
Hee-Soo Choi, Priyansh Trivedi, Mathieu Constant, Karen Fort and Bruno Guillaume	2329
<i>Beyond Static Evaluation: A Dynamic Approach to Assessing AI Assistants' API Invocation Capabilities</i>	
Honglin Mu, Yang Xu, Yunlong Feng, Xiaofeng Han, Yitong Li, Yutai Hou and Wanxiang Che	2342
<i>Beyond the Known: Investigating LLMs Performance on Out-of-Domain Intent Detection</i>	
Pei Wang, Keqing He, Yejie Wang, Xiaoshuai Song, Yutao Mou, Jingang Wang, Yunsen Xian, Xunliang Cai and Weiran Xu	2354
<i>Beyond Words: Decoding Facial Expression Dynamics in Motivational Interviewing</i>	
Nezih Younsi, Catherine Pelachaud and Laurence Chaby	2365
<i>BigNLI: Native Language Identification with Big Bird Embeddings</i>	
Sergey Kramp, Giovanni Cassani and Chris Emmery	2375
<i>Biomedical Concept Normalization over Nested Entities with Partial UMLS Terminology in Russian</i>	
Natalia Loukachevitch, Andrey Sakhovskiy and Elena Tutubalina	2383
<i>Biomedical Entity Linking as Multiple Choice Question Answering</i>	
Zhenxi Lin, Ziheng Zhang, Xian Wu and Yefeng Zheng	2390
<i>Bits and Pieces: Investigating the Effects of Subwords in Multi-task Parsing across Languages and Domains</i>	
Daniel Dakota and Sandra Kübler	2397
<i>BiVert: Bidirectional Vocabulary Evaluation Using Relations for Machine Translation</i>	
Carinne Cherf and Yuval Pinter	2410
<i>BKEE: Pioneering Event Extraction in the Vietnamese Language</i>	
Thi-Nhung Nguyen, Bang Tien Tran, Trong-Nghia Luu, Thien Huu Nguyen and Kiem-Hieu Nguyen	2421
<i>BlendX: Complex Multi-Intent Detection with Blended Patterns</i>	
Yejin Yoon, Jungyeon Lee, Kangsan Kim, Chanhee Park and Taeuk Kim	2428
<i>BLN600: A Parallel Corpus of Machine/Human Transcribed Nineteenth Century Newspaper Texts</i>	
Callum William Booth, Alan Thomas and Robert Gaizauskas	2440
<i>Bootstrapping UMR Annotations for Arapaho from Language Documentation Resources</i>	
Matthew J. Buchholz, Julia Bonn, Claire Benet Post, Andrew Cowell and Alexis Palmer	2447
<i>BootTOD: Bootstrap Task-oriented Dialogue Representations by Aligning Diverse Responses</i>	
Weihao Zeng, Keqing He, Yejie Wang, Dayuan Fu and Weiran Xu	2458

<i>Born a BabyNet with Hierarchical Parental Supervision for End-to-End Text Image Machine Translation</i>	
Cong Ma, Yaping Zhang, Zhiyang Zhang, Yupu Liang, Yang Zhao, Yu Zhou and Chengqing Zong	2468
<i>BP4ER: Bootstrap Prompting for Explicit Reasoning in Medical Dialogue Generation</i>	
Yuhong He, Yongqi Zhang, Shizhu He and Jun Wan	2480
<i>Breakthrough from Nuance and Inconsistency: Enhancing Multimodal Sarcasm Detection with Context-Aware Self-Attention Fusion and Word Weight Calculation.</i>	
Hongfei Xue, Linyan Xu, Yu Tong, Rui Li, Jiali Lin and Dazhi Jiang	2493
<i>Bridging Computational Lexicography and Corpus Linguistics: A Query Extension for OntoLex-FrAC</i>	
Christian Chiarcos, Ranka Stanković, Maxim Ionov and Gilles Sérasset	2504
<i>Bridging Textual and Tabular Worlds for Fact Verification: A Lightweight, Attention-Based Model</i>	
Shirin Dabbaghi Varnosfaderani, Canasai Kruengkrai, Ramin Yahyapour and Junichi Yamagishi	2515
<i>Bridging the Code Gap: A Joint Learning Framework across Medical Coding Systems</i>	
Geunyeong Jeong, Seokwon Jeong, Juoh Sun and Harksoo Kim	2520
<i>Bring Invariant to Variant: A Contrastive Prompt-based Framework for Temporal Knowledge Graph Forecasting</i>	
Ying Zhang, Xinying Qian, Yu Zhao, Baohang Zhou, Kehui Song and Xiaojie Yuan ..	2526
<i>Building a Broad Infrastructure for Uniform Meaning Representations</i>	
Julia Bonn, Matthew J. Buchholz, Jayeol Chun, Andrew Cowell, William Croft, Lukas Denk, Sijia Ge, Jan Hajič, Kenneth Lai, James H. Martin, Skatje Myers, Alexis Palmer, Martha Palmer, Claire Benet Post, James Pustejovsky, Kristine Stenzel, Haibo Sun, Zdeňka Urešová, Rosa Vallejos, Jens E. L. Van Gysel, Meagan Vigus, Nianwen Xue and Jin Zhao	2537
<i>Building a Database of Conversational Routines</i>	
Polina Bychkova, Alyaxey Yaskevich, Serafima Gyulasaryan and Ekaterina Rakhilina	2548
<i>Building a Data Infrastructure for a Mid-Resource Language: The Case of Catalan</i>	
Aitor Gonzalez-Agirre, Montserrat Marimon, Carlos Rodriguez-Penagos, Javier Aula-Blasco, Irene Baucells, Carme Armentano-Oller, Jorge Palomar-Giner, Baybars Kulebi and Marta Villegas	2556
<i>Building a Japanese Document-Level Relation Extraction Dataset Assisted by Cross-Lingual Transfer</i>	
Youmi Ma, An Wang and Naoaki Okazaki	2567
<i>Building MUSCLE, a Dataset for MULTilingual Semantic Classification of Links between Entities</i>	
Lucia Pitarch, Carlos Bobed Lisbona, David Abián, Jorge Gracia and Jordi Bernad ..	2580
<i>Building Question-Answer Data Using Web Register Identification</i>	
Anni Eskelinen, Amanda Myntti, Erik Henriksson, Sampo Pyysalo and Veronika Laippala	2595

<i>CAGK: Collaborative Aspect Graph Enhanced Knowledge-based Recommendation</i> Xiaotong Song, Huiping Lin, Jiatao Zhu and Xinyi Gong	2612
<i>CALAMR: Component ALignment for Abstract Meaning Representation</i> Paul Landes and Barbara Di Eugenio	2622
<i>Calibrating LLM-Based Evaluator</i> Yuxuan Liu, Tianchi Yang, Shaohan Huang, Zihan Zhang, Haizhen Huang, Furu Wei, Weiwei Deng, Feng Sun and Qi Zhang	2638
<i>CAM 2.0: End-to-End Open Domain Comparative Question Answering System</i> ahmad shallouf, Hanna Herasimchyk, Mikhail Salnikov, Rudy Alexandro Garrido Veliz, Natia Mestvirishvili, Alexander Panchenko, Chris Biemann and Irina Nikishina	2657
<i>CAMAL: A Novel Dataset for Multi-label Conversational Argument Move Analysis</i> Viet Dac Lai, Duy Ngoc Pham, Jonathan Steinberg, Jamie Mikeska and Thien Huu Nguyen	2673
<i>Camel Morph MSA: A Large-Scale Open-Source Morphological Analyzer for Modern Standard Arabic</i> Christian Khairallah, Salam Khalifa, Reham Marzouk, Mayar Nassar and Nizar Habash	2683
<i>CamemBERT-bio: Leveraging Continual Pre-training for Cost-Effective Models on French Biomedical Data</i> Rian Touchent and Éric de la Clergerie	2692
<i>CAMERA³: An Evaluation Dataset for Controllable Ad Text Generation in Japanese</i> Go Inoue, Akihiko Kato, Masato Mita, Ukyo Honda and Peinan Zhang	2702
<i>Can Factual Statements Be Deceptive? The DeFaBel Corpus of Belief-based Deception</i> Aswathy Velutharambath, Roman Klinger and Amelie Wüthrl	2708
<i>Can GPT-4 Identify Propaganda? Annotation and Detection of Propaganda Spans in News Articles</i> Maram Hasanain, Fatema Ahmad and Firoj Alam	2724
<i>Can Humans Identify Domains?</i> Maria Barrett, Max Müller-Eberstein, Elisa Bassignana, Amalie Brogaard Pauli, Mike Zhang and Rob van der Goot	2745
<i>Can Language Models Learn Embeddings of Propositional Logic Assertions?</i> Nurul Fajrin Ariyani, Zied Bouraoui, Richard Booth and Steven Schockaert	2766
<i>Can Large Language Models Automatically Score Proficiency of Written Essays?</i> Watheq Ahmad Mansour, Salam Albatarni, Sohaila Eltanbouly and Tamer Elsayed ..	2777
<i>Can Large Language Models Discern Evidence for Scientific Hypotheses? Case Studies in the Social Sciences</i> Sai Koneru, Jian Wu and Sarah Rajtmajer	2787
<i>Can Large Language Models Learn Translation Robustness from Noisy-Source In-context Demonstrations?</i> Leiyu Pan, Yongqi Leng and Deyi Xiong	2798

<i>Can Machine Translation Bridge Multilingual Pretraining and Cross-lingual Transfer Learning?</i> Shaoxiong Ji, Timothee Mickus, Vincent Segonne and Jörg Tiedemann.....	2809
<i>Can Multiple-choice Questions Really Be Useful in Detecting the Abilities of LLMs?</i> Wangyue Li, Liangzhi Li, Tong Xiang, Xiao Liu, Wei Deng and Noa Garcia	2819
<i>Can Small Language Models Help Large Language Models Reason Better?: LM-Guided Chain-of-Thought</i> Jooyoung Lee, Fan Yang, Thanh Tran, Qian Hu, Emre Barut and Kai-Wei Chang	2835
<i>Can We Identify Stance without Target Arguments? A Study for Rumour Stance Classification</i> Yue Li and Carolina Scarton	2844
<i>Can We Learn Question, Answer, and Distractors All from an Image? A New Task for Multiple-choice Visual Question Answering</i> Wenjian Ding, Yao Zhang, Jun Wang, Adam Jatowt and Zhenglu Yang.....	2852
<i>CARE: Co-Attention Network for Joint Entity and Relation Extraction</i> Wenjun Kong and Yamei Xia	2864
<i>CareCorpus: A Corpus of Real-World Solution-Focused Caregiver Strategies for Personalized Pediatric Rehabilitation Service Design</i> Mina Valizadeh, Vera C. Kaelin, Mary A. Khetani and Natalie Parde.....	2871
<i>CASIMIR: A Corpus of Scientific Articles Enhanced with Multiple Author-Integrated Revisions</i> Léane Isabelle Jourdan, Florian Boudin, Nicolas Hernandez and Richard Dufour	2883
<i>Categorical Grammar Induction with Stochastic Category Selection</i> Christian Clark and William Schuler.....	2893
<i>Causal Intersectionality and Dual Form of Gradient Descent for Multimodal Analysis: A Case Study on Hateful Memes</i> Yosuke Miyanishi and Minh Le Nguyen	2901
<i>CBBQ: A Chinese Bias Benchmark Dataset Curated with Human-AI Collaboration for Large Language Models</i> Yufei Huang and Deyi Xiong	2917
<i>CBT-LLM: A Chinese Large Language Model for Cognitive Behavioral Therapy-based Mental Health Question Answering</i> Hongbin Na	2930
<i>CB-Whisper: Contextual Biasing Whisper Using Open-Vocabulary Keyword-Spotting</i> Yuang Li, Yinglu Li, Min Zhang, Chang Su, Jiawei Yu, Mengyao Piao, Xiaosong Qiao, Miaomiao Ma, Yanqing Zhao and Hao Yang	2941
<i>CEPT: A Contrast-Enhanced Prompt-Tuning Framework for Emotion Recognition in Conversation</i> Qingqing Gao, Jiuxin Cao, Biwei Cao, Xin Guan and Bo Liu.....	2947
<i>CE-VDG: Counterfactual Entropy-based Bias Reduction for Video-grounded Dialogue Generation</i> Hongcheng Liu, Pingjie Wang, Zhiyuan Zhu, Yanfeng Wang and Yu Wang	2958

<i>ChainLM: Empowering Large Language Models with Improved Chain-of-Thought Prompting</i> Xiaoxue Cheng, Junyi Li, Wayne Xin Zhao and Ji-Rong Wen	2969
<i>ChainNet: Structured Metaphor and Metonymy in WordNet</i> Rowan Hall Maudslay, Simone Teufel, Francis Bond and James Pustejovsky	2984
<i>Challenges in Pre-Training Graph Neural Networks for Context-Based Fake News Detection: An Evaluation of Current Strategies and Resource Limitations</i> Gregor Donabauer and Udo Kruschwitz	2997
<i>Challenging Negative Gender Stereotypes: A Study on the Effectiveness of Automated Counter-Stereotypes</i> Isar Nejadgholi, Kathleen C. Fraser, Anna Kerkhof and Svetlana Kiritchenko	3005
<i>Characteristic AI Agents via Large Language Models</i> Xi Wang, Hongliang Dai, Shen Gao and Piji Li	3016
<i>Character-level Language Models for Abbreviation and Long-form Detection</i> Leonardo Zilio, Shenbin Qian, Diptesh Kanojia and Constantin Orasan	3028
<i>Charles Translator: A Machine Translation System between Ukrainian and Czech</i> Martin Popel, Lucie Polakova, Michal Novák, Jindřich Helcl, Jindřich Libovický, Pavel Straňák, Tomas Krabac, Jaroslava Hlavacova, Mariia Anisimova and Tereza Chlanova	3038
<i>Charting the Linguistic Landscape of Developing Writers: An Annotation Scheme for Enhancing Native Language Proficiency</i> Miguel Da Corte and Jorge Baptista	3046
<i>ChartThinker: A Contextual Chain-of-Thought Approach to Optimized Chart Summarization</i> Mengsha Liu, Daoyuan Chen, Yaliang Li, Guian Fang and Ying Shen	3057
<i>ChatASU: Evoking LLM's Reflexion to Truly Understand Aspect Sentiment in Dialogues</i> Yiding Liu, Jingjing Wang, Jiamin Luo, Tao Zeng and Guodong Zhou	3075
<i>ChatEL: Entity Linking with Chatbots</i> Yifan Ding, Qingkai Zeng and Tim Weninger	3086
<i>ChatGPT Is a Knowledgeable but Inexperienced Solver: An Investigation of Commonsense Problem in Large Language Models</i> Ning Bian, Xianpei Han, Le Sun, Hongyu Lin, Yaojie Lu, Ben He, Shanshan Jiang and Bin Dong	3098
<i>ChatGPT Rates Natural Language Explanation Quality like Humans: But on Which Scales?</i> Fan Huang, Haewoon Kwak, Kunwoo Park and Jisun An	3111
<i>ChatGPT Role-play Dataset: Analysis of User Motives and Model Naturalness</i> Yufei Tao, Ameeta Agrawal, Judit Dombi, Tetyana Sydorenko and Jung In Lee	3133
<i>ChatUIE: Exploring Chat-based Unified Information Extraction Using Large Language Models</i> Jun Xu, Mengshu Sun, Zhiqiang Zhang and Jun Zhou	3146

<i>CHICA: A Developmental Corpus of Child-Caregiver's Face-to-face vs. Video Call Conversations in Middle Childhood</i>	
Dhia Elhak Goumri, Abhishek Agrawal, Mitja Nikolaus, Hong Duc Thang Vu, Kübra Bodur, Elias Emmar, Cassandre Armand, Chiara Mazzocconi, Shreejata Gupta, Laurent Prévot, Benoit Favre, Leonor Becerra-Bonache and Abdellah Fourtassi	3153
<i>Chinese Morpheme-informed Evaluation of Large Language Models</i>	
Yaqi Yin, Yue Wang and Yang Liu	3165
<i>Chinese Sequence Labeling with Semi-Supervised Boundary-Aware Language Model Pre-training</i>	
Longhui Zhang, Dingkun Long, Meishan Zhang, Yanzhao Zhang, Pengjun Xie and Min Zhang	3179
<i>CHisIEC: An Information Extraction Corpus for Ancient Chinese History</i>	
Xuemei Tang, Qi Su, Jun Wang and Zekun Deng	3192
<i>Chitchat as Interference: Adding User Backstories to Task-Oriented Dialogues</i>	
Armand Stricker and Patrick Paroubek	3203
<i>Choice-75: A Dataset on Decision Branching in Script Learning</i>	
Zhaoyi Hou, Li Zhang and Chris Callison-Burch	3215
<i>C-Journal: A Journaling Application for Detecting and Classifying Cognitive Distortions Using Deep-Learning Based on a Crowd-sourced Dataset</i>	
Nada Elsharawi and Alia El Bolock	3224
<i>Claim-Centric and Sentiment Guided Graph Attention Network for Rumour Detection</i>	
Sajad Ramezani, Mauzama Firdaus and Lili Mou	3235
<i>Clarifying Completions: Evaluating How LLMs Respond to Incomplete Questions</i>	
Angus Adlesee, Oliver Lemon and Arash Eshghi	3242
<i>Classifying Social Media Users before and after Depression Diagnosis via Their Language Usage: A Dataset and Study</i>	
Falwah Alhamed, Julia Ive and Lucia Specia	3250
<i>Class-Incremental Few-Shot Event Detection</i>	
Kailin Zhao, Xiaolong Jin, Long Bai, Jiafeng Guo and Xueqi Cheng	3261
<i>CLASSLA-web: Comparable Web Corpora of South Slavic Languages Enriched with Linguistic and Genre Annotation</i>	
Nikola Ljubešić and Taja Kuzman	3271
<i>CLAUSE-ATLAS: A Corpus of Narrative Information to Scale up Computational Literary Analysis</i>	
Enrica Troiano and Piek T.J.M. Vossen	3283
<i>CLEVR-POC: Reasoning-Intensive Visual Question Answering in Partially Observable Environments</i>	
Savitha Sam Abraham, Marjan Alirezaie and Luc de Raedt	3297
<i>CLFFRD: Curriculum Learning and Fine-grained Fusion for Multimodal Rumor Detection</i>	
Fan Xu, Lei Zeng, Bowei Zou, Ai Ti Aw and Huan Rong	3314

<i>CLHA: A Simple Yet Effective Contrastive Learning Framework for Human Alignment</i> Feiteng Fang, Liang Zhu, Xi Feng, Jinchang Hou, Qixuan Zhao, Chengming Li, Xiping Hu, Ruifeng Xu and Min Yang	3325
<i>CLiCK: A Benchmark Dataset of Cultural and Linguistic Intelligence in Korean</i> Eunsu Kim, Juyoung Suk, Philhoon Oh, Haneul Yoo, James Thorne and Alice Oh ...	3335
<i>Clue-Instruct: Text-Based Clue Generation for Educational Crossword Puzzles</i> Andrea Zugarini, Kamyar Zeinalipour, Surya Sai Kadali, Marco Maggini, Marco Gori and Leonardo Rigutini	3347
<i>CMDAG: A Chinese Metaphor Dataset with Annotated Grounds as CoT for Boosting Metaphor Generation</i> Yujie Shao, Xinrong Yao, Xingwei Qu, Chenghua Lin, Shi Wang, Wenhao Huang, Ge Zhang and Jie Fu	3357
<i>CMNEE: A Large-Scale Document-Level Event Extraction Dataset Based on Open-Source Chi- nese Military News</i> Mengna Zhu, Zijie Xu, Kaisheng Zeng, Kaiming Xiao, Mao Wang, Wenjun Ke and Hongbin Huang	3367
<i>CM-Off-Meme: Code-Mixed Hindi-English Offensive Meme Detection with Multi-Task Learning by Leveraging Contextual Knowledge</i> Gitanjali Kumari, Dibyanayan Bandyopadhyay, Asif Ekbal and Vinutha B. NarayanaMurthy	3380
<i>CO3: Low-resource Contrastive Co-training for Generative Conversational Query Rewrite</i> Yifei Yuan, Chen Shi, Wang Runze, Liyi Chen, Renjun Hu, Zengming Zhang, Feijun Jiang and Wai Lam	3394
<i>CoANZSE Audio: Creation of an Online Corpus for Linguistic and Phonetic Analysis of Australian and New Zealand Englishes</i> Steven Coats	3407
<i>Coarse-Tuning for Ad-hoc Document Retrieval Using Pre-trained Language Models</i> Atsushi Keyaki and Ribeka Keyaki	3413
<i>CoBaLD Annotation: The Enrichment of the Enhanced Universal Dependencies with the Se- mantical Pattern</i> Maria Andreevna Petrova, Alexandra M. Ivoylova and Anastasia Tishchenkova	3422
<i>CoCoMIC: Code Completion by Jointly Modeling In-file and Cross-file Context</i> Yangruibo Ding, Zijian Wang, Wasi U. Ahmad, Murali Krishna Ramanathan, Ramesh Nal- lapati, Parminder Bhatia, Dan Roth and Bing Xiang	3433
<i>Code Defect Detection Using Pre-trained Language Models with Encoder-Decoder via Line- Level Defect Localization</i> Jimin An, YunSeok Choi and Jee-Hyong Lee	3446
<i>Code-Mixed Probes Show How Pre-Trained Models Generalise on Code-Switched Text</i> Frances Adriana Laureano De Leon, Harish Tayyar Madabushi and Mark Lee	3457

<i>Code-Mixed Text Augmentation for Latvian ASR</i> Martins Kronis, Askars Salimbajevs and Mārcis Pinnis	3469
<i>Cognitive Information Bottleneck: Extracting Minimal Sufficient Cognitive Language Processing Signals</i> Yuto Harada and Yohei Oseki	3480
<i>CollabKG: A Learnable Human-Machine-Cooperative Information Extraction Toolkit for (Event) Knowledge Graph Construction</i> Xiang Wei, Yufeng Chen, Ning Cheng, Xingyu Cui, Jinan Xu and Wenjuan Han	3490
<i>Collecting and Analyzing Dialogues in a Tagline Co-Writing Task</i> Xulin Zhou, Takuma Ichikawa and Ryuichiro Higashinaka	3507
<i>Collecting Human-Agent Dialogue Dataset with Frontal Brain Signal toward Capturing Unexpressed Sentiment</i> Shun Katada, Ryu Takeda and Kazunori Komatani	3518
<i>Collecting Linguistic Resources for Assessing Children's Pronunciation of Nordic Languages</i> Anne Marte Haug Olstad, Anna Smolander, Sofia Strömbergsson, Sari Ylinen, Minna Lehtonen, Mikko Kurimo, Yaroslav Getman, Tamás Grósz, Xinwei Cao, Torbjørn Svendsen and Giampiero Salvi	3529
<i>Combining Discourse Coherence with Large Language Models for More Inclusive, Equitable, and Robust Task-Oriented Dialogue</i> Katherine Atwell, Mert Inan, Anthony B. Sicilia and Malihe Alikhani	3538
<i>COMET for Low-Resource Machine Translation Evaluation: A Case Study of English-Maltese and Spanish-Basque</i> Júlia Falcão, Claudia Borg, Nora Aranberri and Kurt Abela	3553
<i>COMICORDA: Dialogue Act Recognition in Comic Books</i> Jiri Martinek, Pavel Kral, Ladislav Lenc and Josef Baloun	3566
<i>Common European Language Data Space</i> Georg Rehm, Stelios Piperidis, Khalid Choukri, Andrejs Vasiljevs, Katrin Marheinecke, Victoria Arranz, Aivars Bērziņš, Miltos Deligiannis, Dimitris Galanis, Maria Giagkou, Katerina Gkirtzou, Dimitris Gkoumas, Annika Grützner-Zahn, Athanasia Kolovou, Penny Labropoulou, Andis Lagzdīņš, Elena Leitner, Valérie Mapelli, Hélène Mazo, Simon Ostermann, Stefania Racioppa, Mickaël Rigault and Leon Voukoutis	3579
<i>Common Ground Tracking in Multimodal Dialogue</i> Ibrahim Khalil Khebour, Kenneth Lai, Mariah Bradford, Yifan Zhu, Richard A. Brutti, Christopher Tam, Jingxuan Tu, Benjamin A. Ibarra, Nathaniel Blanchard, Nikhil Krishnaswamy and James Pustejovsky	3587
<i>Comparative Analysis of Sign Language Interpreting Agents Perception: A Study of the Deaf</i> Alfarabi Imashev, Nurziya Oralbayeva, Gulmira Baizhanova and Anara Sandygulova	3603
<i>Comparing Static and Contextual Distributional Semantic Models on Intrinsic Tasks: An Evaluation on Mandarin Chinese Datasets</i> A Pranav, Yan Cong, Emmanuele Chersoni, Yu-Yin Hsu and Alessandro Lenci	3610

<i>Comparison of Conventional Hybrid and CTC/Attention Decoders for Continuous Visual Speech Recognition</i>	
David Gimeno-Gómez and Carlos-D. Martínez-Hinarejos	3628
<i>Comparison of the Intimacy Process between Real and Acting-based Long-term Text Chats</i>	
Tsunehiro Arimoto, Hiroaki Sugiyama, Hiromi Narimatsu and Masahiro Mizukami ...	3639
<i>Complex Word Identification: A Comparative Study between ChatGPT and a Dedicated Model for This Task</i>	
Abdelhak Keliou, Mathieu Constant and Christophe Coeur	3645
<i>Comprehensive Study on German Language Models for Clinical and Biomedical Text Understanding</i>	
Ahmad Idrissi-Yaghir, Amin Dada, Henning Schäfer, Kamyar Arzideh, Giulia Baldini, Jan Trienes, Max Hasin, Jeanette Bewersdorff, Cynthia S. Schmidt, Marie Bauer, Kaleb E. Smith, Jiang Bian, Yonghui Wu, Jörg Schlötterer, Torsten Zesch, Peter A. Horn, Christin Seifert, Felix Nensa, Jens Kleesiek and Christoph M. Friedrich	3654
<i>Computational Modelling of Plurality and Definiteness in Chinese Noun Phrases</i>	
Yuqi Liu, Guanyi Chen and Kees van Deemter	3666
<i>CONAN-MT-SP: A Spanish Corpus for Counternarrative Using GPT Models</i>	
María Estrella Vallecillo Rodríguez, María Victoria Cantero Romero, Isabel Cabrera De Castro, Arturo Montejo Ráez and María Teresa Martín Valdivia	3677
<i>Conceptual Pacts for Reference Resolution Using Small, Dynamically Constructed Language Models: A Study in Puzzle Building Dialogues</i>	
Julian Hough, Sina Zarriß, Casey Kennington, David Schlangen and Massimo Poesio	3689
<i>ConEC: Earnings Call Dataset with Real-world Contexts for Benchmarking Contextual Speech Recognition</i>	
Ruizhe Huang, Mahsa Yarmohammadi, Jan Trmal, Jing Liu, Desh Raj, Leibny Paola Garcia, Alexei V. Ivanov, Patrick Ehlen, Mingzhi Yu, Dan Povey and Sanjeev Khudanpur	3700
<i>Conjoin after Decompose: Improving Few-Shot Performance of Named Entity Recognition</i>	
Chengcheng Han, Renyu Zhu, Jun Kuang, Fengjiao Chen, Xiang Li, Ming Gao, Xuezhi Cao and Yunsen Xian	3707
<i>CoNLL#: Fine-grained Error Analysis and a Corrected Test Set for CoNLL-03 English</i>	
Andrew Rueda, Elena Alvarez-Mellado and Constantine Lignos	3718
<i>Connecting Language Technologies with Rich, Diverse Data Sources Covering Thousands of Languages</i>	
Daan van Esch, Sandy Ritchie, Sebastian Ruder, Julia Kreutzer, Clara Rivera, Ishank Saxena and Isaac Caswell	3729
<i>Constructing a Dependency Treebank for Second Language Learners of Korean</i>	
Hakyung Sung and Gyu-Ho Shin	3747
<i>Constructing Indonesian-English Travelogue Dataset</i>	
Eunike Andriani Kardinata, Hiroki Ouchi and Taro Watanabe	3759

<i>Constructing Korean Learners' L2 Speech Corpus of Seven Languages for Automatic Pronunciation Assessment</i>	
Seunghye Han, Sunhee Kim and Minhwa Chung	3772
<i>Construction of Paired Knowledge Graph - Text Datasets Informed by Cyclic Evaluation</i>	
Ali Mousavi, Xin Zhan, He Bai, Peng Shi, Theodoros Rekatsinas, Benjamin Han, Yunyao Li, Jeffrey Pound, Joshua M. Susskind, Natalie Schluter, Ihab F. Ilyas and Navdeep Jaitly	3782
<i>Constructions Are So Difficult That Even Large Language Models Get Them Right for the Wrong Reasons</i>	
Shijia Zhou, Leonie Weissweiler, Taiqi He, Hinrich Schütze, David R. Mortensen and Lori Levin.....	3804
<i>Context-Aware Non-Autoregressive Document-Level Translation with Sentence-Aligned Connectionist Temporal Classification</i>	
Hao Yu, Kaiyu Huang, Anqi Zhao, Junpeng Liu and Degen Huang	3812
<i>Context Matters: Enhancing Metaphor Recognition in Proverbs</i>	
Gamze Goren and Carlo Strapparava.....	3825
<i>Context Shapes Emergent Communication about Concepts at Different Levels of Abstraction</i>	
Kristina Kobrock, Xenia Isabel Ohmer, Elia Bruni and Nicole Gotzner.....	3831
<i>Contextualizing Generated Citation Texts</i>	
Biswadip Mandal, Xiangci Li and Jessica Ouyang	3849
<i>Contextual Modeling for Document-level ASR Error Correction</i>	
Jin Jiang, Xunjian Yin, Xiaojun Wan, Wei Peng, Rongjun Li, Jingyuan Yang and Yanquan Zhou.....	3855
<i>Continual Few-shot Event Detection via Hierarchical Augmentation Networks</i>	
Chenlong Zhang, Pengfei Cao, Yubo Chen, Kang Liu, Zhiqiang Zhang, Mengshu Sun and Jun Zhao	3868
<i>Continual Reinforcement Learning for Controlled Text Generation</i>	
Velizar Shulev and Khalil Sima'an	3881
<i>Continued Pre-training on Sentence Analogies for Translation with Small Data</i>	
Liyang Wang, Haotong Wang and Yves Lepage.....	3890
<i>Continuous Relational Diffusion Driven Topic Model with Multi-grained Text for Microblog</i>	
Chenhao Wu, Ruifang He, Chang Liu and Bo Wang	3897
<i>ContrastWSD: Enhancing Metaphor Detection with Word Sense Disambiguation Following the Metaphor Identification Procedure</i>	
Mohamad MZ Elzohbi and Richard Zhao	3907
<i>Contribution of Move Structure to Automatic Genre Identification: An Annotated Corpus of French Tourism Websites</i>	
Rémi Cardon, Trang Tran Hanh Pham, Julien Zakhia Doueihi and Thomas François.	3916
<i>Controllable Paraphrase Generation for Semantic and Lexical Similarities</i>	
Yuya Ogasa, Tomoyuki Kajiwara and Yuki Arase	3927

<i>Controllable Sentence Simplification in Swedish Using Control Prefixes and Mined Paraphrases</i> Julius Monsen and Arne Jonsson	3943
<i>Controlled Generation with Prompt Insertion for Natural Language Explanations in Grammatical Error Correction</i> Masahiro Kaneko and Naoaki Okazaki	3955
<i>ControversialQA: Exploring Controversy in Question Answering</i> Zhen Wang, Peide Zhu and Jie Yang	3962
<i>Conversational Grounding: Annotation and Analysis of Grounding Acts and Grounding Units</i> Biswesh Mohapatra, Seemab Hassan, Laurent Romary and Justine Cassell	3967
<i>Converting Legacy Data to CLDF: A FAIR Exit Strategy for Linguistic Web Apps</i> Robert Forkel, Daniel G. Swanson and Steven Moran	3978
<i>CookingSense: A Culinary Knowledgebase with Multidisciplinary Assertions</i> Donghee Choi, Mogan Gim, Donghyeon Park, Mujeen Sung, Hyunjae Kim, Jaewoo Kang and Jihun Choi	3983
<i>CoRelation: Boosting Automatic ICD Coding through Contextualized Code Relation Learning</i> Junyu Luo, Xiaochen Wang, Jiaqi Wang, Aofei Chang, Yaqing Wang and Fenglong Ma	3997
<i>CORI: CJKV Benchmark with Romanization Integration - a Step towards Cross-lingual Transfer beyond Textual Scripts</i> Hoang Nguyen, Chenwei Zhang, Ye Liu, Natalie Parde, Eugene Rohrbaugh and Philip S. Yu	4008
<i>Corpus Creation and Automatic Alignment of Historical Dutch Dialect Speech</i> Martijn Bentum, Eric Sanders, Antal P.J. van den Bosch, Douwe Zeldenrust and Henk van den Heuvel	4021
<i>Corpus Services: A Framework to Curate XML Corpus Data</i> Aleksandr Riaposov and Elena Lazarenko	4030
<i>Correcting Language Model Bias for Text Classification in True Zero-Shot Learning</i> Feng Zhao, Wan Xianlin, Cheng Yan and Chu Kiong Loo	4036
<i>Correcting Pronoun Homophones with Subtle Semantics in Chinese Speech Recognition</i> Zhaobo Zhang, Rui Gan, Pingpeng Yuan and Hai Jin	4047
<i>Correlations between Multilingual Language Model Geometry and Crosslingual Transfer Performance</i> Cheril Shah, Yashashree Chandak, Atharv Mahesh Mane, Benjamin Bergen and Tyler A. Chang	4059
<i>Cost-Effective Discourse Annotation in the Prague Czech–English Dependency Treebank</i> Jiří Mírovský, Pavlína Synková, Lucie Polakova and Marie Paclíková	4067
<i>Counterfactual Dialog Mixing as Data Augmentation for Task-Oriented Dialog Systems</i> Sebastian Steindl, Ulrich Schäfer and Bernd Ludwig	4078

<i>Creating Terminological Resources in the Digital Age for Less-resourced Languages</i> Mercè Vázquez	4088
<i>Creation and Analysis of an International Corpus of Privacy Laws</i> Sonu Gupta, Geetika Gopi, Harish Balaji, Ellen Poplavska, Nora O’Toole, Siddhant Arora, Thomas Norton, Norman Sadeh and Shomir Wilson	4092
<i>Croatian Idioms Integration: Enhancing the Lldioms Multilingual Linked Idioms Dataset</i> Ivana Filipović Petrović, Miguel López Otal and Slobodan Beliga	4106
<i>CroCoSum: A Benchmark Dataset for Cross-Lingual Code-Switched Summarization</i> Ruo Chen Zhang and Carsten Eickhoff	4113
<i>Cross-Lingual Learning vs. Low-Resource Fine-Tuning: A Case Study with Fact-Checking in Turkish</i> Recep Firat Cekineli, Çağrı Çöltekin and Pinar Karagoz	4127
<i>Cross-lingual Named Entity Corpus for Slavic Languages</i> Jakub Piskorski, Michał Marcińczuk and Roman Yangarber	4143
<i>Cross-Lingual NLU: Mitigating Language-Specific Impact in Embeddings Leveraging Adversarial Learning</i> SaedeH Tahery, Sahar Kianian and Saeed Farzi	4158
<i>Cross-lingual Transfer or Machine Translation? On Data Augmentation for Monolingual Semantic Textual Similarity</i> Sho Hoshino, Akihiko Kato, Soichiro Murakami and Peinan Zhang	4164
<i>Cross-Lingual Transfer Robustness to Lower-Resource Languages on Adversarial Datasets</i> Shadi Manafi and Nikhil Krishnaswamy	4174
<i>CrossTune: Black-Box Few-Shot Classification with Label Enhancement</i> Danqing Luo, Chen Zhang, Yan Zhang and Haizhou Li	4185
<i>Cross-type French Multiword Expression Identification with Pre-trained Masked Language Models</i> Van-Tuan Bui and Agata Savary	4198
<i>CSSWiki: A Chinese Sentence Simplification Dataset with Linguistic and Content Operations</i> Fengkai Liu and John S. Y. Lee	4205
<i>CTSM: Combining Trait and State Emotions for Empathetic Response Model</i> Yufeng Wang, Chao Chen, Zhou Yang, Shuhui Wang and Xiangwen Liao	4214
<i>CulturaX: A Cleaned, Enormous, and Multilingual Dataset for Large Language Models in 167 Languages</i> Thuat Nguyen, Chien Van Nguyen, Viet Dac Lai, Hieu Man, Nghia Trung Ngo, Franck Deroncourt, Ryan A. Rossi and Thien Huu Nguyen	4226
<i>Curation of Benchmark Templates for Measuring Gender Bias in Named Entity Recognition Models</i> Ana Cimitan, Ana Alves Pinto and Michaela Geierhos	4238

<i>CuRIAM: Corpus Re Interpretation and Metalanguage in U.S. Supreme Court Opinions</i> Michael Kranzlein, Nathan Schneider and Kevin Tobia	4247
<i>Curriculum Learning Meets Directed Acyclic Graph for Multimodal Emotion Recognition</i> Cam-Van Thi Nguyen, Cao-Bach Nguyen, Duc-Trong Le and Quang-Thuy Ha	4259
<i>CuSINeS: Curriculum-driven Structure Induced Negative Sampling for Statutory Article Retrieval</i> Santosh T.Y.S.S., Kristina Kaiser and Matthias Grabmair	4266
<i>CWTM: Leveraging Contextualized Word Embeddings from BERT for Neural Topic Modeling</i> Zheng Fang, Yulan He and Rob Procter	4273
<i>CyberAgressionAdo-v2: Leveraging Pragmatic-Level Information to Decipher Online Hate in French Multiparty Chats</i> Anais Ollagnier	4287
<i>Czech Dataset for Complex Aspect-Based Sentiment Analysis Tasks</i> Jakub Šmíd, Pavel Přibáň, Ondrej Prazak and Pavel Kral	4299
<i>DACL: Disfluency Augmented Curriculum Learning for Fluent Text Generation</i> Rohan Chaudhury, Maria Teleki, Xiangjue Dong and James Caverlee	4311
<i>DADIT: A Dataset for Demographic Classification of Italian Twitter Users and a Comparison of Prediction Methods</i> Lorenzo Lupo, Paul Bose, Mahyar Habibi, Dirk Hovy and Carlo Schwarz	4322
<i>DANCER: Entity Description Augmented Named Entity Corrector for Automatic Speech Recognition</i> Yi-Cheng Wang, Hsin-Wei Wang, Bi-Cheng Yan, Chi-Han Lin and Berlin Chen	4333
<i>DanteLLM: Let's Push Italian LLM Research Forward!</i> Andrea Bacciu, Cesare Campagnano, Giovanni Trappolini and Fabrizio Silvestri	4343
<i>DARIUS: A Comprehensive Learner Corpus for Argument Mining in German-Language Essays</i> Nils-Jonathan Schaller, Andrea Horbach, Lars Ingver Höft, Yuning Ding, Jan Luca Bahr, Jennifer Meyer and Thorben Jansen	4356
<i>Data Collection Pipeline for Low-Resource Languages: A Case Study on Constructing a Tetun Text Corpus</i> Gabriel de Jesus and Sérgio Sobral Nunes	4368
<i>Data Drift in Clinical Outcome Prediction from Admission Notes</i> Paul Grundmann, Jens-Michalis Papaioannou, Tom Oberhauser, Thomas Steffek, Amy Siu, Wolfgang Nejdl and Alexander Loeser	4381
<i>Data-Informed Global Sparseness in Attention Mechanisms for Deep Neural Networks</i> Ileana Rugina, Rumen Dangovski, Li Jing, Preslav Nakov and Marin Soljacic	4392
<i>Dataset for Identification of Homophobia and Transphobia for Telugu, Kannada, and Gujarati</i> Prasanna Kumar Kumaresan, Rahul Ponnusamy, Dhruv Sharma, Paul Buitelaar and Bharathi Raja Chakravarthi	4404
<i>Dataset of Quotation Attribution in German News Articles</i> Fynn Petersen-Frey and Chris Biemann	4412

<i>DC-MBR: Distributional Cooling for Minimum Bayesian Risk Decoding</i> Jianhao Yan, Jin Xu, Fandong Meng, Jie Zhou and Yue Zhang	4423
<i>DDxGym: Online Transformer Policies in a Knowledge Graph Based Natural Language Environment</i> Benjamin Winter, Alexei Gustavo Figueroa Rosero, Alexander Loeser, Felix Alexander Gers, Nancy Katerina Figueroa Rosero and Ralf Krestel	4438
<i>Dealing with Data Scarcity in Spoken Question Answering</i> Merve Ünlü Menevşe, Yusufcan Manav, Ebru Arisoy and Arzucan Özgür	4449
<i>Debiasing Multi-Entity Aspect-Based Sentiment Analysis with Norm-Based Data Augmentation</i> Scott Friedman, Joan Zheng and Hillel Steinmetz	4456
<i>Deciphering Emotional Landscapes in the Iliad: A Novel French-Annotated Dataset for Emotion Recognition</i> Davide Picca and John Pavlopoulos	4462
<i>DECM: Evaluating Bilingual ASR Performance on a Code-switching/mixing Benchmark</i> Enes Yavuz Ugan, Ngoc-Quan Pham and Alexander Waibel	4468
<i>Decoding at the Speed of Thought: Harnessing Parallel Decoding of Lexical Units for LLMs</i> Chenxi Sun, Hongzhi Zhang, Zijia Lin, Jingyuan Zhang, Fuzheng Zhang, Zhongyuan Wang, Bin Chen, Chengru Song, Di Zhang, Kun Gai and Deyi Xiong	4476
<i>Decoding Probing: Revealing Internal Linguistic Structures in Neural Language Models Using Minimal Pairs</i> Linyang He, Peili Chen, Ercong Nie, Yuanning Li and Jonathan R. Brennan	4488
<i>Decompose, Prioritize, and Eliminate: Dynamically Integrating Diverse Representations for Multimodal Named Entity Recognition</i> Zihao Zheng, Zihan Zhang, Zexin Wang, Ruiji Fu, Ming Liu, Zhongyuan Wang and Bing Qin	4498
<i>Deconstructing In-Context Learning: Understanding Prompts via Corruption</i> Namrata Shivagunde, Vladislav Lialin, Sherin Muckatira and Anna Rumshisky	4509
<i>DEEM: Dynamic Experienced Expert Modeling for Stance Detection</i> Xiaolong Wang, Yile Wang, Sijie Cheng, Peng Li and Yang Liu	4530
<i>Deep Learning Based Named Entity Recognition Models for Recipes</i> Ayush Agarwal, Janak Kapuriya, Shubham Agrawal, Akhil Vamshi Konam, Mansi Goel, Rishabh Gupta, Shrey Rastogi, Niharika Niharika and Ganesh Bagler	4542
<i>Deep Reinforcement Learning-based Dialogue Policy with Graph Convolutional Q-network</i> Kai Xu, Zhengyu Wang, Yuxuan Long and Qiaona Zhao	4555
<i>Deep Reinforcement Learning with Hierarchical Action Exploration for Dialogue Generation</i> Itsubun Cho, Ryota Takahashi, Yusaku Yanase and Hiroaki Saito	4566
<i>DeFaktS: A German Dataset for Fine-Grained Disinformation Detection through Social Media Framing</i> Shaina Ashraf, Isabel Bezzaoui, Ionut Andone, Alexander Markowetz, Jonas Fegert and Lucie Flek	4580

<i>DEIE: Benchmarking Document-level Event Information Extraction with a Large-scale Chinese News Dataset</i>	
Yubing Ren, Yanan Cao, Hao Li, Yingjie Li, Zixuan Ma, Fang Fang, Ping Guo and Wei Ma	4592
<i>DELAN: Dual-Level Alignment for Vision-and-Language Navigation by Cross-Modal Contrastive Learning</i>	
Mengfei Du, Binhao Wu, Jiwen Zhang, Zhihao Fan, Zejun Li, Ruipu Luo, Xuanjing Huang and Zhongyu Wei	4605
<i>Demonstration Retrieval-Augmented Generative Event Argument Extraction</i>	
Shiming He, Yu Hong, Shuai Yang, Jianmin Yao and Guodong Zhou	4617
<i>Denoising Labeled Data for Comment Moderation Using Active Learning</i>	
Andraž Pelicon, Mladen Karan, Ravi Shekhar, Matthew Purver and Senja Pollak	4626
<i>Denoising Table-Text Retrieval for Open-Domain Question Answering</i>	
Deokhyung Kang, Baikjin Jung, Yunsu Kim and Gary Geunbae Lee	4634
<i>Dependencies over Times and Tools (DoTT)</i>	
Andy Luecking, Giuseppe Abrami, Leon Hammerla, Marc Rahn, Daniel Baumartz, Steffen Eger and Alexander Mehler	4641
<i>Depth Aware Hierarchical Replay Continual Learning for Knowledge Based Question Answering</i>	
Zhixiong Cao, Hai-Tao Zheng, Yangning Li, Jin Xu, Rongsheng Li and Hong-Gee Kim	4654
<i>Depth-Wise Attention (DWAtt): A Layer Fusion Method for Data-Efficient Classification</i>	
Muhammad ElNokrashy, Badr AlKhamissi and mona Diab	4665
<i>Deriving Entity-Specific Embeddings from Multi-Entity Sequences</i>	
Connor Heaton and Prasenjit Mitra	4675
<i>DET: A Dual-Encoding Transformer for Relational Graph Embedding</i>	
Lingbing Guo, Zhuo Chen, Jiaoyan Chen, Qiang Zhang and Huajun Chen	4685
<i>Detecting Conceptual Abstraction in LLMs</i>	
Michaela Regneri, Alhassan Abdelhalim and Soeren Laue	4697
<i>Detecting Critical Errors Considering Cross-Cultural Factors in English-Korean Translation</i>	
Sugyeong Eo, Jungwoo Lim, Chanjun Park, DaHyun Jung, Seonmin Koo, Hyeonseok Moon, Jaehyung Seo and Heuseok Lim	4705
<i>Detecting Cybercrimes in Accordance with Pakistani Law: Dataset and Evaluation Using PLMs</i>	
Faizad Ullah, Ali Faheem, Ubaid Azam, Muhammad Sohaib Ayub, Faisal Kamiran and Asim Karim	4717
<i>Detecting Hallucination and Coverage Errors in Retrieval Augmented Generation for Controversial Topics</i>	
Tyler A. Chang, Katrin Tomanek, Jessica Hoffmann, Nithum Thain, Erin MacMurray van Liemt, Kathleen Meier-Hellstern and Lucas Dixon	4729
<i>Detecting Impact Relevant Sections in Scientific Research</i>	
Maria Becker, Kanyao Han, Antonina Werthmann, Rezvaneh Rezapour, Haejin Lee and Jana Diesner	4744

<i>Detecting Loanwords in Emakhuwa: An Extremely Low-Resource Bantu Language Exhibiting Significant Borrowing from Portuguese</i>	
Felermimo Dario Mario Ali, Henrique Lopes Cardoso and Rui Sousa-Silva	4750
<i>Detecting Offensive Language in an Open Chatbot Platform</i>	
Hyeonho Song, Jisu Hong, Chani Jung, Hyojin Chin, Mingi Shin, Yubin Choi, Junghoi Choi and Meeyoung Cha	4760
<i>Detecting Sexual Content at the Sentence Level in First Millennium Latin Texts</i>	
Thibault Clerice	4772
<i>Detection, Diagnosis, and Explanation: A Benchmark for Chinese Medial Hallucination Evaluation</i>	
Chengfeng Dou, Ying Zhang, Yanyuan Chen, Zhi Jin, Wenpin Jiao, Haiyan Zhao and Yu Huang	4784
<i>Developing a Benchmark for Pronunciation Feedback: Creation of a Phonemically Annotated Speech Corpus of isiZulu Language Learner Speech</i>	
Alexandra O’Neil, Nils Hjortnaes, Francis Tyers, Zinhle Nkosi, Thulile Ndlovu, Zanele Mlondo and Ngami Phumzile Pewa	4795
<i>Developing a Rhetorical Structure Theory Treebank for Czech</i>	
Lucie Polakova, Jiří Mírovský, Šárka Zikánová and Eva Hajicova	4802
<i>Development and Evaluation of Pre-trained Language Models for Historical Danish and Norwegian Literary Texts</i>	
Ali Al-Laith, Alexander Conroy, Jens Bjerring-Hansen and Daniel Hershcovich	4811
<i>Development of Community-Oriented Text-to-Speech Models for Māori ‘Avaiki Nui (Cook Islands Māori)</i>	
Jesin James, Rolando Coto-Solano, Sally Akevai Nicholas, Joshua Zhu, Bovey Yu, Fuki Babasaki, Jenny Tyler Wang and Nicholas Derby	4820
<i>DGoT: Dynamic Graph of Thoughts for Scientific Abstract Generation</i>	
Xinyu Ning, Yutong Zhao, Yitong Liu and Hongwen Yang	4832
<i>DGS-Fabeln-1: A Multi-Angle Parallel Corpus of Fairy Tales between German Sign Language and German Text</i>	
Fabrizio Nunnari, Eleftherios Avramidis, Cristina España-Bonet, Marco González, Anna Hennes and Patrick Gebhard	4847
<i>Dialogue Systems Can Generate Appropriate Responses without the Use of Question Marks?— a Study of the Effects of “?” for Spoken Dialogue Systems –</i>	
Tomoya Mizumoto, Takato Yamazaki, Katsumasa Yoshikawa, Masaya Ohagi, Toshiki Kawamoto and Toshinori Sato	4858
<i>DiaSet: An Annotated Dataset of Arabic Conversations</i>	
Abraham Israeli, Aviv Naaman, Guy Maduel, Rawaa Makhoul, Dana Qaraeen, Amir Ejmail, Dina Lisnanskey, Julian Jubran, Shai Fine and Kfir Bar	4865
<i>Did You Get It? A Zero-Shot Approach to Locate Information Transfers in Conversations</i>	
Eliot Maës, Hossam Boudraa, Philippe Blache and Leonor Becerra-Bonache	4877

<i>Difficulty-Focused Contrastive Learning for Knowledge Tracing with a Large Language Model-Based Difficulty Prediction</i>	
Unggi Lee, Sungjun Yoon, Joon Seo Yun, Kyoungsoo Park, YoungHoon Jung, Damji Stratton and Hyeoncheol Kim	4891
<i>Diffusion Based Counterfactual Augmentation for Dual Sentiment Classification</i>	
Dancheng Xin, Jiawei Yuan and Yang Li	4901
<i>DiffusionDialog: A Diffusion Model for Diverse Dialog Generation with Latent Space</i>	
Jianxiang Xiang, Zhenhua Liu, Haodong Liu, Yin Bai, Jia Cheng and Wenliang Chen	4912
<i>DimA: A Parameter-efficient Fine-tuning Method with Knowledge Transfer Based on Transformer</i>	
Wenxuan Zhang, Min Huang, Zhuoyang Song and Qinghai Miao	4922
<i>Disambiguating Homographs and Homophones Simultaneously: A Regrouping Method for Japanese</i>	
Yo Sato	4935
<i>DiscoGeM 2.0: A Parallel Corpus of English, German, French and Czech Implicit Discourse Relations</i>	
Frances Yung, Merel Scholman, Sarka Zikanova and Vera Demberg	4940
<i>Discourse Structure for the Minecraft Corpus</i>	
Kate Thompson, Julie Hunter and Nicholas Asher	4957
<i>Discriminative Language Model as Semantic Consistency Scorer for Prompt-based Few-Shot Text Classification</i>	
Zhipeng Xie and Yahe Li	4968
<i>Disentangling Pretrained Representation to Leverage Low-Resource Languages in Multilingual Machine Translation</i>	
Frederikus Hudi, Zhi Qu, Hidetaka Kamigaito and Taro Watanabe	4978
<i>DISRPT: A Multilingual, Multi-domain, Cross-framework Benchmark for Discourse Processing</i>	
Chloé Braud, Amir Zeldes, Laura Rivière, Yang Janet Liu, Philippe Muller, Damien Sileo and Tatsuya Aoyama	4990
<i>Distantly Supervised Contrastive Learning for Low-Resource Scripting Language Summarization</i>	
Junzhe Liang, Haifeng Sun, Zirui Zhuang, Qi Qi, Jingyu Wang and Jianxin Liao	5006
<i>Distillation with Explanations from Large Language Models</i>	
Hanyu Zhang, Xiting Wang, Xiang Ao and Qing He	5018
<i>Distill, Fuse, Pre-train: Towards Effective Event Causality Identification with Commonsense-Aware Pre-trained Model</i>	
Peixin Huang, Xiang Zhao, Minghao Hu, Zhen Tan and Weidong Xiao	5029
<i>Distilling Causal Effect of Data in Continual Few-shot Relation Learning</i>	
Weihang Ye, Peng Zhang, Jing Zhang, Hui Gao and Moyao Wang	5041
<i>Distractor Generation Using Generative and Discriminative Capabilities of Transformer-based Models</i>	
Shiva Taslimipoor, Luca Benedetto, Mariano Felice and Paula Buttery	5052

<i>Distribution Aware Metrics for Conditional Natural Language Generation</i> David M. Chan, Yiming Ni, David Ross, Sudheendra Vijayanarasimhan, Austin Myers and John Canny	5064
<i>Diversifying Question Generation over Knowledge Base via External Natural Questions</i> Shasha Guo, Jing Zhang, Xirui Ke, Cuiping Li and Hong Chen	5096
<i>DMON: A Simple Yet Effective Approach for Argument Structure Learning</i> Sun Wei, Mingxiao Li, Jingyuan Sun, Jesse Davis and Marie-Francine Moens	5109
<i>Doc2SoarGraph: Discrete Reasoning over Visually-Rich Table-Text Documents via Semantic-Oriented Hierarchical Graphs</i> Fengbin Zhu, Chao Wang, Fuli Feng, Zifeng Ren, Moxin Li and Tat-Seng Chua	5119
<i>DOC-RAG: ASR Language Model Personalization with Domain-Distributed Co-occurrence Retrieval Augmentation</i> Puneet Mathur, Zhe Liu, Ke Li, Yingyi Ma, Gil Karen, Zeeshan Ahmed, Dinesh Manocha and Xuedong Zhang	5132
<i>DocScript: Document-level Script Event Prediction</i> Puneet Mathur, Vlad I. Morariu, Aparna Garimella, Franck Dernoncourt, Jiuxiang Gu, Ramit Sawhney, Preslav Nakov, Dinesh Manocha and Rajiv Jain	5140
<i>Document-Level Event Extraction via Information Interaction Based on Event Relation and Argument Correlation</i> Bangze Pan, Yang Li, Suge Wang, Xiaoli Li, Deyu Li, Jian Liao and Jianxing Zheng .	5156
<i>Document Set Expansion with Positive-Unlabeled Learning Using Intractable Density Estimation</i> Haiyang Zhang, Qiuyi Chen, Yanjie Zou, Jia Wang, Yushan Pan and Mark Stevenson	5167
<i>Do Emergent Abilities Exist in Quantized Large Language Models: An Empirical Study</i> Peiyu Liu, Zikang Liu, Ze-Feng Gao, Dawei Gao, Wayne Xin Zhao, Yaliang Li, Bolin Ding and Ji-Rong Wen	5174
<i>Does ChatGPT Know That It Does Not Know? Evaluating the Black-Box Calibration of ChatGPT</i> Youliang Yuan, Wenxuan Wang, Qingshuo Guo, Yiming Xiong, Chihao Shen and Pinjia He	5191
<i>Does the Generator Mind Its Contexts? An Analysis of Generative Model Faithfulness under Context Transfer</i> Xinshuo Hu, Dongfang Li, Xiaoguang Li, Yuxiang Wu, Lifeng Shang and Baotian Hu	5202
<i>Does the Language Matter? Curriculum Learning over Neo-Latin Languages</i> Giulia Pucci and Leonardo Ranaldi	5212
<i>Do Language Models Care about Text Quality? Evaluating Web-Crawled Corpora across 11 Languages</i> Rik van Noord, Taja Kuzman, Peter Rupnik, Nikola Ljubešić, Miquel Esplà-Gomis, Gema Ramírez-Sánchez and Antonio Toral	5221
<i>Do Large Language Models Understand Mansplaining? Well, Actually...</i> Carla Perez Almendros and Jose Camacho-Collados	5235

<i>Domain Adaptation for Dense Retrieval and Conversational Dense Retrieval through Self-Supervision by Meticulous Pseudo-Relevance Labeling</i>	
Minghan Li and Eric Gaussier	5247
<i>Domain-Agnostic Adapter Architecture for Deception Detection: Extensive Evaluations with the DIFraud Benchmark</i>	
Dainis A. Bumber, Fatima Zahra Qachfar and Rakesh Verma	5260
<i>Domain-aware and Co-adaptive Feature Transformation for Domain Adaption Few-shot Relation Extraction</i>	
Yijun Liu, Feifei Dai, Xiaoyan Gu, Minghui Zhai, Bo Li and Meiou Zhang.....	5275
<i>Domain Generalization via Causal Adjustment for Cross-Domain Sentiment Analysis</i>	
Siyin Wang, Jie Zhou, Qin Chen, Qi Zhang, Tao Gui and Xuanjing Huang	5286
<i>Domain Transferable Semantic Frames for Expert Interview Dialogues</i>	
Taishi Chika, Taro Okahisa, Takashi Kodama, Yin Jou Huang, Yugo Murawaki and Sadao Kurohashi.....	5299
<i>Do Neural Language Models Inferentially Compose Concepts the Way Humans Can?</i>	
Amilleah Rodriguez, Shaonan Wang and Liina Pykkänen	5309
<i>DORE: A Dataset for Portuguese Definition Generation</i>	
Anna Beatriz Dimas Furtado, Tharindu Ranasinghe, Frederic Blain and Ruslan Mitkov.....	5315
<i>DOSA: A Dataset of Social Artifacts from Different Indian Geographical Subcultures</i>	
Agrima Seth, Sanchit Ahuja, Kalika Bali and Sunayana Sitaram	5323
<i>DP-CRE: Continual Relation Extraction via Decoupled Contrastive Learning and Memory Structure Preservation</i>	
Mengyi Huang, Meng Xiao, Ludi Wang and Yi Du	5338
<i>Dr3: Ask Large Language Models Not to Give Off-Topic Answers in Open Domain Multi-Hop Question Answering</i>	
Yuan Gao, Yiheng Zhu, Yuanbin Cao, Yinzhi Zhou, Zhen Wu, Yujie Chen, Shenglan Wu, Haoyuan Hu and Xinyu Dai.....	5350
<i>DRAMA: Dynamic Multi-Granularity Graph Estimate Retrieval over Tabular and Textual Question Answering</i>	
Ruize Yuan, Xiang Ao, Li Zeng and Qing He	5365
<i>DrBenchmark: A Large Language Understanding Evaluation Benchmark for French Biomedical Domain</i>	
Yanis Labrak, Adrien Bazoge, Oumaima El Khettari, Mickael Rouvier, pacome constant dit beaufils, Natalia Grabar, Béatrice Daille, Solen Quiniou, Emmanuel Morin, Pierre-Antoine Gourraud and Richard Dufour	5376
<i>Dual Complex Number Knowledge Graph Embeddings</i>	
Yao Dong, Qingchao Kong, Lei Wang and Yin Luo	5391
<i>Dual Encoder: Exploiting the Potential of Syntactic and Semantic for Aspect Sentiment Triplet Extraction</i>	
Xiaowei Zhao, Yong Zhou and Xiujuan Xu	5401

<i>DuetSim: Building User Simulator with Dual Large Language Models for Task-Oriented Dialogues</i>	
Xiang Luo, Zhiwen Tang, Jin Wang and Xuejie Zhang	5414
<i>Dynamic Knowledge Prompt for Chest X-ray Report Generation</i>	
Shenshen Bu, Yujie Song, Taiji Li and Zhiming Dai	5425
<i>Dynamic Reward Adjustment in Multi-Reward Reinforcement Learning for Counselor Reflection Generation</i>	
Do June Min, Veronica Perez-Rosas, Ken Resnicow and Rada Mihalcea	5437
<i>Dynamic Spatial-Temporal Aggregation for Skeleton-Aware Sign Language Recognition</i>	
Lianyu Hu, Liqing Gao, Zekang Liu and Wei Feng	5450
<i>EcoVerse: An Annotated Twitter Dataset for Eco-Relevance Classification, Environmental Impact Analysis, and Stance Detection</i>	
Francesca Grasso, Stefano Locci, Giovanni Siragusa and Luigi Di Caro	5461
<i>ECtHR-PCR: A Dataset for Precedent Understanding and Prior Case Retrieval in the European Court of Human Rights</i>	
Santosh T.Y.S.S., Rashid Haddad and Matthias Grabmair	5473
<i>EDDA: An Encoder-Decoder Data Augmentation Framework for Zero-Shot Stance Detection</i>	
Daijun Ding, Li Dong, Zhichao Huang, Guangning Xu, Xu Huang, Bo Liu, Liwen Jing and Bowen Zhang	5484
<i>EDEN: A Dataset for Event Detection in Norwegian News</i>	
Samia Touileb, Jeanett Murstad, Petter Mæhlum, Lubos Steskal, Lilja Charlotte Storset, Huiling You and Lilja Øvrelid	5495
<i>Educational Dialogue Systems for Visually Impaired Students: Introducing a Task-Oriented User-Agent Corpus</i>	
Elisa Di Nuovo, Manuela Sanguinetti, Pier Felice Balestrucci, Luca Anselma, Cristian Bernareggi and Alessandro Mazzei	5507
<i>EEE-QA: Exploring Effective and Efficient Question-Answer Representations</i>	
Zhanghao Hu, Yijun Yang, Junjie Xu, Yifu Qiu and Pinzhen Chen	5520
<i>Eesthetic: A Paralex Lexicon of Estonian Paradigms</i>	
Sacha Beniamine, Mari Aigro, Matthew Baerman, Jules Bouton and Maria Copot ...	5526
<i>Effective Distillation of Table-based Reasoning Ability from LLMs</i>	
Bohao Yang, Chen Tang, Kun Zhao, Chenghao Xiao and Chenghua Lin	5538
<i>Effective Integration of Text Diffusion and Pre-Trained Language Models with Linguistic Easy-First Schedule</i>	
Yimin Ou and Ping Jian	5551
<i>Efficiency and Effectiveness in Task-Oriented Dialogue: On Construction Repetition, Information Rate, and Task Success</i>	
Jun Sen Yee, Mario Giulianelli and Arabella J. Sinclair	5562
<i>Efficient AMR Parsing with CLAP: Compact Linearization with an Adaptable Parser</i>	
Abelardo Carlos Martinez Lorenzo and Roberto Navigli	5578

<i>Efficient and Accurate Contextual Re-Ranking for Knowledge Graph Question Answering</i> Kexuan Sun, Nicolaas Paul Jedema, Karishma Sharma, Ruben Janssen, Jay Pujara, Pedro Szekely and Alessandro Moschitti	5585
<i>EFTNAS: Searching for Efficient Language Models in First-Order Weight-Reordered Super-Networks</i> Juan Pablo Munoz, Yi Zheng and Nilesh Jain	5596
<i>Eliciting Motivational Interviewing Skill Codes in Psychotherapy with LLMs: A Bilingual Dataset and Analytical Study</i> Xin Sun, Jiahuan Pei, Jan de Wit, Mohammad Aliannejadi, Emiel Kraemer, Jos T.P. Dobber and Jos A. Bosch	5609
<i>ELLEN: Extremely Lightly Supervised Learning for Efficient Named Entity Recognition</i> Haris Riaz, Razvan Gabriel Dumitru and Mihai Surdeanu	5622
<i>EMAD: A Bridge Tagset for Unifying Arabic POS Annotations</i> Omar Kallas, Go Inoue and Nizar Habash	5637
<i>Emancipating Event Extraction from the Constraints of Long-Tailed Distribution Data Utilizing Large Language Models</i> Zhigang Kan, Liwen Peng, Linbo Qiao and Dongsheng Li	5644
<i>EMOLIS App and Dataset to Find Emotionally Close Cartoons</i> Soëlie Lerch, Patrice Bellot, Elisabeth Murisasco and Emmanuel Bruno	5654
<i>EmoProgress: Cumulated Emotion Progression Analysis in Dreams and Customer Service Dialogues</i> Eileen Wemmer, Sofie Labat and Roman Klinger	5660
<i>EmoPrompt-ECPE: Emotion Knowledge-aware Prompt-tuning for Emotion-Cause Pair Extraction</i> Xue Gu, Zhihan Zhou, Ziyao Meng, Jian Li, Tiago Gomes, Adriano Tavares and Hao Xu	5678
<i>Emotags: Computer-Assisted Verbal Labelling of Expressive Audiovisual Utterances for Expressive Multimodal TTS</i> Gérard Bailly, Romain Legrand, Martin Lenglet, Frédéric Elisei, Maëva Hueber and Olivier Perrotin	5689
<i>Emotion Analysis in NLP: Trends, Gaps and Roadmap for Future Directions</i> Flor Miriam Plaza-del-Arco, Alba A. Cercas Curry, Amanda Cercas Curry and Dirk Hovy	5696
<i>Emotion Recognition in Conversation via Dynamic Personality</i> Yan Wang, Bo Wang, Yachao Zhao, Dongming Zhao, Xiaojia Jin, Jijun Zhang, Ruifang He and Yuexian Hou	5711
<i>EmoTrans: Emotional Transition-based Model for Emotion Recognition in Conversation</i> Zhongquan Jian, Ante Wang, Jinsong Su, Junfeng Yao, Meihong Wang and Qingqiang Wu	5723

<i>EmpCRL: Controllable Empathetic Response Generation via In-Context Commonsense Reasoning and Reinforcement Learning</i>	
Mingxiu Cai, Daling Wang, Shi Feng and Yifei Zhang	5734
<i>Empowering Low-Resource Regional Languages with Lexicons : A Comparative Study of NLP Tools for Morphosyntactic Analysis</i>	
Cristina Garcia Holgado and Marianne Vergez-Couret	5747
<i>Empowering Oneida Language Revitalization: Development of an Oneida Verb Conjugator</i>	
Yanfei Lu, Patrick Littell and Keren Rice.....	5757
<i>Empowering Small-Scale Knowledge Graphs: A Strategy of Leveraging General-Purpose Knowledge Graphs for Enriched Embeddings</i>	
Albert Sawczyn, Jakub Binkowski, Piotr Bielak and Tomasz Kajdanowicz.....	5768
<i>Empowering Tree-structured Entailment Reasoning: Rhetorical Perception and LLM-driven Interpretability</i>	
Longyin Zhang, Bowei Zou and Ai Ti Aw	5783
<i>Emstremo: Adapting Emotional Support Response with Enhanced Emotion-Strategy Integrated Selection</i>	
Junlin Li, Bo Peng and Yu-Yin Hsu.....	5794
<i>Encoding Gesture in Multimodal Dialogue: Creating a Corpus of Multimodal AMR</i>	
Kenneth Lai, Richard Brutti, Lucia Donatelli and James Pustejovsky.....	5806
<i>Endowing Neural Language Learners with Human-like Biases: A Case Study on Dependency Length Minimization</i>	
Yuqing Zhang, Tessa Verhoef, Gertjan van Noord and Arianna Bisazza	5819
<i>End-to-end Parsing of Procedural Text into Flow Graphs</i>	
Dhaivat J. Bhatt, Seyed Ahmad Abdollahpouri Hosseini, Federico Fancellu and Afsaneh Fazly	5833
<i>Enhanced Coherence-Aware Network with Hierarchical Disentanglement for Aspect-Category Sentiment Analysis</i>	
Jin Cui, Fumiyo Fukumoto, Xinfeng Wang, Yoshimi Suzuki, Jiyi Li, Noriko Tomuro and Wanzeng Kong.....	5843
<i>Enhanced Facet Generation with LLM Editing</i>	
Joosung Lee and Jinhong Kim.....	5856
<i>Enhance Robustness of Language Models against Variation Attack through Graph Integration</i>	
Zi Xiong, Lizhi Qing, Yangyang Kang, Jiawei Liu, Hongsong Li, Changlong Sun, Xiaozhong Liu and Wei Lu.....	5866
<i>Enhancing Code Generation Performance of Smaller Models by Distilling the Reasoning Ability of LLMs</i>	
Zhihong Sun, Chen Lyu, Bolun Li, Yao Wan, Hongyu Zhang, Ge Li and Zhi Jin	5878
<i>Enhancing Court View Generation with Knowledge Injection and Guidance</i>	
Ang Li, Yiquan Wu, Yifei Liu, Kun Kuang, Fei Wu and Ming Cai	5896

<i>Enhancing Cross-Document Event Coreference Resolution by Discourse Structure and Semantic Information</i>	
Qiang Gao, Bobo Li, Zixiang Meng, Yunlong Li, Jun Zhou, Fei Li, Chong Teng and Donghong Ji	5907
<i>Enhancing Distantly Supervised Named Entity Recognition with Strong Label Guided Lottery Training</i>	
Zhiyuan Ma, Jintao Du, Changhua Meng and Weiqiang Wang	5922
<i>Enhancing Effectiveness and Robustness in a Low-Resource Regime via Decision-Boundary-aware Data Augmentation</i>	
Kyohoon Jin, Junho Lee, Juhwan Choi, Sangmin Song and Youngbin Kim	5930
<i>Enhancing Emotion Prediction in News Headlines: Insights from ChatGPT and Seq2Seq Models for Free-Text Generation</i>	
Ge Gao, Jongin Kim, Sejin Paik, Ekaterina Novozhilova, Yi Liu, Sarah T. Bonna, Margrit Betke and Derry Tanti Wijaya	5944
<i>Enhancing Few-Shot Topic Classification with Verbalizers. a Study on Automatic Verbalizer and Ensemble Methods</i>	
Quang Anh Nguyen, Nadi Tomeh, Mustapha Lebbah, Thierry Charnois, Hanene Azzag and Santiago Cordoba Muñoz	5956
<i>Enhancing Hindi Feature Representation through Fusion of Dual-Script Word Embeddings</i>	
Lianxi Wang, Yujia Tian and Zhuowei Chen	5966
<i>Enhancing Image-to-Text Generation in Radiology Reports through Cross-modal Multi-Task Learning</i>	
Nurbanu Aksoy, Nishant Ravikumar and Serge Sharoff	5977
<i>Enhancing Knowledge Retrieval with Topic Modeling for Knowledge-Grounded Dialogue</i>	
Nhat Tran and Diane Litman	5986
<i>Enhancing Knowledge Selection via Multi-level Document Semantic Graph</i>	
Haoran Zhang and Tan Yongmei	5996
<i>Enhancing Large Language Models through Transforming Reasoning Problems into Classification Tasks</i>	
Tarun Raheja, Raunak Sinha, Advit Deepak, Will Healy, Jayanth Srinivasa, Myungjin Lee and Ramana Kompella	6007
<i>Enhancing Low-Resource LLMs Classification with PEFT and Synthetic Data</i>	
Parth Patwa, Simone Filice, Zhiyu Chen, Giuseppe Castellucci, Oleg Rokhlenko and Shervin Malmasi	6017
<i>Enhancing Parameter-efficient Fine-tuning with Simple Calibration Based on Stable Rank</i>	
Peiyu Liu, Ze-Feng Gao, Xiao Zhang, Wayne Xin Zhao and Ji-Rong Wen	6024
<i>Enhancing Phrase Representation by Information Bottleneck Guided Text Diffusion Process for Keyphrase Extraction</i>	
Yuanzhen Luo, Qingyu Zhou and Feng Zhou	6036

<i>Enhancing Scientific Document Summarization with Research Community Perspective and Background Knowledge</i>	
Sudipta Singha Roy and Robert E. Mercer	6048
<i>Enhancing Semantics in Multimodal Chain of Thought via Soft Negative Sampling</i>	
Guangmin Zheng, Jin Wang, Xiaobing Zhou and Xuejie Zhang	6059
<i>Enhancing Taiwanese Hokkien Dual Translation by Exploring and Standardizing of Four Writing Systems</i>	
Bo-Han Lu, Yi-Hsuan Lin, Annie Lee and Richard Tzong-Han Tsai	6077
<i>Enhancing Text-to-SQL Capabilities of Large Language Models through Tailored Promptings</i>	
Zhao Tan, Xiping Liu, Qing Shu, Xi Li, Changxuan Wan, Dexi Liu, Qizhi Wan and Guoqiong Liao	6091
<i>Enhancing Translation Ability of Large Language Models by Leveraging Task-Related Layers</i>	
Pei Cheng, Xiayang Shi and Yinlin Li	6110
<i>Enhancing Unrestricted Cross-Document Event Coreference with Graph Reconstruction Networks</i>	
Loic de Langhe, Orphee de Clercq and Veronique Hoste	6122
<i>Enhancing Writing Proficiency Classification in Developmental Education: The Quest for Accuracy</i>	
Miguel Da Corte and Jorge Baptista	6134
<i>Enhancing Zero-Shot Chain-of-Thought Reasoning in Large Language Models through Logic</i>	
Xufeng Zhao, Mengdi Li, Wenhao Lu, Cornelius Weber, Jae Hee Lee, Kun Chu and Stefan Wermter	6144
<i>Enough Is Enough! a Case Study on the Effect of Data Size for Evaluation Using Universal Dependencies</i>	
Rob van der Goot, Zoey Liu and Max Müller-Eberstein	6167
<i>Enriching a Time-Domain Astrophysics Corpus with Named Entity, Coreference and Astrophysical Relationship Annotations</i>	
Atilla Kaan Alkan, Felix Grezes, Cyril Grouin, Fabian Schussler and Pierre Zweigenbaum	6177
<i>Enriching Word Usage Graphs with Cluster Definitions</i>	
Andrey Kutuzov, Mariia Fedorova, Dominik Schlechtweg and Nikolay Arefyev	6189
<i>Ensembles of Hybrid and End-to-End Speech Recognition.</i>	
Aditya Kamlesh Parikh, Louis ten Bosch and Henk van den Heuvel	6199
<i>EpiGEN: An Efficient Multi-API Code GENeration Framework under Enterprise Scenario</i>	
Sijie Li, Sha Li, Hao Zhang, Shuyang Li, Kai Chen, Jianyong Yuan, Yi Cao and Lvqing Yang	6206
<i>EpLSA: Synergy of Expert-prefix Mixtures and Task-Oriented Latent Space Adaptation for Diverse Generative Reasoning</i>	
Fujun Zhang, Xiangdong Su, Jiang Li, Rong Yan and Guanglai Gao	6216

<i>EPOQUE: An English-Persian Quality Estimation Dataset</i> Mohammed Hossein Jafari Harandi, Fatemeh Azadi, Mohammad Javad Dousti and Hesham Faili.....	6228
<i>EROS:Entity-Driven Controlled Policy Document Summarization</i> Joykirat Singh, Sehban Fazili, Rohan Jain and Md. Shad Akhtar	6236
<i>Error Analysis of NLP Models and Non-Native Speakers of English Identifying Sarcasm in Reddit Comments</i> Oliver Cakebread-Andrews, Le An Ha, Ingo Frommholz and Burcu Can.....	6247
<i>Error-Robust Retrieval for Chinese Spelling Check</i> Xunjian Yin, Xinyu Hu, Jin Jiang and Xiaojun Wan	6257
<i>EsCoLA: Spanish Corpus of Linguistic Acceptability</i> Nuria Bel, Marta Punsola and Valle Ruíz-Fernández	6268
<i>ESCP: Enhancing Emotion Recognition in Conversation with Speech and Contextual Prefixes</i> Xiujuan Xu, Xiaoxiao Shi, Zhehuan Zhao and Yu Liu	6278
<i>ESDM: Early Sensing Depression Model in Social Media Streams</i> Bichen Wang, Yuzhe Zi, Yanyan Zhao, Pengfei Deng and Bing Qin	6288
<i>Esposito: An English-Persian Scientific Parallel Corpus for Machine Translation</i> Mersad Esalati, Mohammad Javad Dousti and Hesham Faili.....	6299
<i>Estimating Lexical Complexity from Document-Level Distributions</i> Sondre Wold, Petter Mæhlum and Oddbjørn Hove	6309
<i>Estimating the Causal Effects of Natural Logic Features in Transformer-Based NLI Models</i> Julia Rozanova, Marco Valentino and André Freitas.....	6319
<i>Ethical Reasoning and Moral Value Alignment of LLMs Depend on the Language We Prompt Them in</i> Utkarsh Agarwal, Kumar Tanmay, Aditi Khandelwal and Monojit Choudhury	6330
<i>EthioLLM: Multilingual Large Language Models for Ethiopian Languages with Task Evaluation</i> Atnafu Lambebo Tonja, Israel Abebe Azime, Tadesse Destaw Belay, Mesay Gameda Yigezu, Moges Ahmed Ah Mehamed, Abinew Ali Ayele, Ebrahim Chekol Jibril, Michael Melese Wold-eyohannis, Olga Kolesnikova, Philipp Slusallek, Dietrich Klakow and Seid Muhie Yimam .	6341
<i>European Language Grid: One Year after</i> Georg Rehm, Stelios Piperidis, Dimitris Galanis, Penny Labropoulou, Maria Giagkou, Miltos Deligiannis, Leon Voukoutis, Martin Courtois, Julian Moreno-Schneider and Katrin Marheinecke	6353
<i>Evaluating Automatic Subtitling: Correlating Post-editing Effort and Automatic Metrics</i> Alina Karakanta, Mauro Cettolo, Matteo Negri and Luisa Bentivogli.....	6363
<i>Evaluating ChatGPT against Functionality Tests for Hate Speech Detection</i> Mithun Das, Saurabh Kumar Pandey and Animesh Mukherjee	6370
<i>Evaluating Code-Switching Translation with Large Language Models</i> Muhammad Huzaifah, Weihua Zheng, Nattapol Chanpaisit and Kui Wu	6381

<i>Evaluating Gender Bias of Pre-trained Language Models in Natural Language Inference by Considering All Labels</i>	
Panatchakorn Anantaprayoon, Masahiro Kaneko and Naoaki Okazaki	6395
<i>Evaluating Generative Language Models in Information Extraction as Subjective Question Correction</i>	
Yuchen Fan, Yantao Liu, Zijun Yao, Jifan Yu, Lei Hou and Juanzi Li	6409
<i>Evaluating Performance of Pre-trained Word Embeddings on Assamese, a Low-resource Language</i>	
Dhrubajyoti Pathak, Sukumar Nandi and Priyankoo Sarmah	6418
<i>Evaluating Prompting Strategies for Grammatical Error Correction Based on Language Proficiency</i>	
Min Zeng, Jiexin Kuang, Mengyang Qiu, Jayoung Song and Jungyeul Park	6426
<i>Evaluating Saliency Explanations in NLP by Crowdsourcing</i>	
Xiaotian Lu, Jiyi Li, Zhen Wan, Xiaofeng Lin, Koh Takeuchi and Hisashi Kashima	6431
<i>Evaluating Self-Supervised Speech Representations for Indigenous American Languages</i>	
Chih-Chen Chen, William Chen, Rodolfo Joel Zevallos and John E. Ortega	6444
<i>Evaluating Shortest Edit Script Methods for Contextual Lemmatization</i>	
Olia Toporkov and Rodrigo Agerri	6451
<i>Evaluating Text-to-Speech Synthesis from a Large Discrete Token-based Speech Language Model</i>	
Siyang Wang and Eva Szekely	6464
<i>Evaluating the Efficacy of Large Acoustic Model for Documenting Non-Orthographic Tribal Languages in India</i>	
Tonmoy Rajkhowa, Amartya Roy Chowdhury, Hrishikesh Ravindra Karande and S. R. Mahadeva Prasanna	6475
<i>Evaluating the IWSLT2023 Speech Translation Tasks: Human Annotations, Automatic Metrics, and Segmentation</i>	
Matthias Sperber, Ondřej Bojar, Barry Haddow, Dávid Javorský, Xutai Ma, Matteo Negri, Jan Niehues, Peter Polák, Elizabeth Salesky, Katsuhito Sudoh and Marco Turchi	6484
<i>Evaluating the Potential of Language-family-specific Generative Models for Low-resource Data Augmentation: A Faroese Case Study</i>	
Barbara Scalvini and Iben Nyholm Debess	6496
<i>Evaluating the Quality of a Corpus Annotation Scheme Using Pretrained Language Models</i>	
Furkan Akkurt, Onur Gungor, Büşra Marşan, Tunga Gungor, Balkiz Ozturk Basaran, Arzuçan Özgür and Susan Uskudarli	6504
<i>Evaluating Topic Model on Asymmetric and Multi-Domain Financial Corpus</i>	
Corentin Masson and Patrick Paroubek	6515
<i>Evaluating Unsupervised Dimensionality Reduction Methods for Pretrained Sentence Embeddings</i>	
Gaifan Zhang, Yi Zhou and Danushka Bollegala	6530

<i>Evaluating Webcam-based Gaze Data as an Alternative for Human Rationale Annotations</i> Stephanie Brandl, Oliver Eberle, Tiago Ribeiro, Anders Søgaard and Nora Hollenstein	6544
<i>Evaluating Word Expansion for Multilingual Sentiment Analysis of Parliamentary Speech</i> Yana Nikolova and Costanza Navarretta	6557
<i>Evaluating Workflows for Creating Orthographic Transcripts for Oral Corpora by Transcribing from Scratch or Correcting ASR-Output</i> Jan Gorisch and Thomas Schmidt	6564
<i>Evaluation Dataset for Lexical Translation Consistency in Chinese-to-English Document-level Translation</i> Xiangyu Lei, Junhui Li, Shimin Tao and Hao Yang	6575
<i>Evaluation of Really Good Grammatical Error Correction</i> Robert Östling, Katarina Gillholm, Murathan Kurfalı, Marie Mattson and Mats Wirén	6582
<i>Event-enhanced Retrieval in Real-time Search</i> Yanan Zhang, Xiaoling Bai and Tianhua Zhou	6594
<i>Event Extraction in Basque: Typologically Motivated Cross-Lingual Transfer-Learning Analysis</i> Mikel Zubillaga, Oscar Sainz, Ainara Estarrona, Oier Lopez de Lacalle and Eneko Agirre	6607
<i>EventGround: Narrative Reasoning by Grounding to Eventuality-centric Knowledge Graphs</i> Cheng Jiayang, Lin Qiu, Chunkit Chan, Xin Liu, Yangqiu Song and Zheng Zhang	6622
<i>Event Representation Learning with Multi-Grained Contrastive Learning and Triple-Mixture of Experts</i> Tianqi Hu, Lishuang Li, Xueyang Qin and Yubo Feng	6643
<i>Every Verb in Its Right Place? A Roadmap for Operationalizing Developmental Stages in the Acquisition of L2 German</i> Josef Ruppenhofer, Matthias Schwendemann, Annette Portmann, Katrin Wisniewski and Torsten Zesch	6655
<i>Evidence-guided Inference for Neutralized Zero-shot Transfer</i> Xiaotong Feng, Meng-Fen Chiang, Wang-Chien Lee and Zixin Kuang	6671
<i>EVil-Probe - a Composite Benchmark for Extensive Visio-Linguistic Probing</i> Marie Bexte, Andrea Horbach and Torsten Zesch	6682
<i>EvoGrad: A Dynamic Take on the Winograd Schema Challenge with Human Adversaries</i> Jing Han Sun and Ali Emami	6701
<i>Evolving Knowledge Distillation with Large Language Models and Active Learning</i> Chengyuan Liu, Fubang Zhao, Kun Kuang, Yangyang Kang, Zhuoren Jiang, Changlong Sun and Fei Wu	6717
<i>Examining Temporalities on Stance Detection towards COVID-19 Vaccination</i> Yida Mu, Mali Jin, Kalina Bontcheva and Xingyi Song	6732
<i>Examining the Limitations of Computational Rumor Detection Models Trained on Static Datasets</i> Yida Mu, Xingyi Song, Kalina Bontcheva and Nikolaos Aletras	6739

<i>Executing Natural Language-Described Algorithms with Large Language Models: An Investigation</i>	
Xin Zheng, Qiming Zhu, Hongyu Lin, Yaojie Lu, Xianpei Han and Le Sun	6752
<i>Experimental versus In-Corpus Variation in Referring Expression Choice</i>	
T. Mark Ellison and Fahime Same	6838
<i>Experiments on Speech Synthesis for Teochew, Can Taiwanese Help ?</i>	
Pierre Magistry, Ilaine Wang and Ty Eng Lim	6849
<i>Explainable Multi-hop Question Generation: An End-to-End Approach without Intermediate Question Labeling</i>	
Seonjeong Hwang, Yunsu Kim and Gary Geunbae Lee	6855
<i>Explaining Pre-Trained Language Models with Attribution Scores: An Analysis in Low-Resource Settings</i>	
Wei Zhou, Heike Adel, Hendrik Schuff and Ngoc Thang Vu	6867
<i>Explicit over Implicit: Explicit Diversity Conditions for Effective Question Answer Generation</i>	
Vikas Yadav, Hyuk joon Kwon, Vijay Srinivasan and Hongxia Jin	6876
<i>Exploring and Mitigating Shortcut Learning for Generative Large Language Models</i>	
Zechen Sun, Yisheng Xiao, Juntao Li, Yixin Ji, Wenliang Chen and Min Zhang	6883
<i>Exploring BERT-Based Classification Models for Detecting Phobia Subtypes: A Novel Tweet Dataset and Comparative Analysis</i>	
Anik Das, Milton King and James Alexander Hughes	6894
<i>Exploring Geometric Representational Disparities between Multilingual and Bilingual Translation Models</i>	
Neha Verma, Kenton Murray and Kevin Duh	6909
<i>Exploring Interpretability of Independent Components of Word Embeddings with Automated Word Intruder Test</i>	
Tomáš Musil and David Mareček	6922
<i>Exploring Neural Topic Modeling on a Classical Latin Corpus</i>	
Ginevra Martinelli, Paola Impicciché, Elisabetta Fersini, Francesco Mambriani and Marco Passarotti	6929
<i>Exploring Pathological Speech Quality Assessment with ASR-Powered Wav2Vec2 in Data-Scarce Context</i>	
Tuan Nguyen, Corinne Fredouille, Alain Ghio, Mathieu Balaguer and Virginie Woisard	6935
<i>Exploring the Emotional Dimension of French Online Toxic Content</i>	
Valentina Dragos, Delphine Battistelli, Fatou Sow and Aline Etienne	6945
<i>Exploring the Generalization of Cancer Clinical Trial Eligibility Classifiers across Diseases</i>	
Yumeng Yang	6955
<i>Exploring the Impact of Human Evaluator Group on Chat-Oriented Dialogue Evaluation</i>	
Sarah E. Finch, James D. Finch and Jinho D. Choi	6966

<i>Exploring the Potential of Large Language Models (LLMs) for Low-resource Languages: A Study on Named-Entity Recognition (NER) and Part-Of-Speech (POS) Tagging for Nepali Language</i> Bipesh Subedi, Sunil Regmi, Bal Krishna Bal and Praveen Acharya.....	6974
<i>Exploring the Synergy of Dual-path Encoder and Alignment Module for Better Graph-to-Text Generation</i> Tianxin Zhao, Yingxin Liu, Xiangdong Su, Jiang Li and Guanglai Gao	6980
<i>Exploring the Usability of Persuasion Techniques for Downstream Misinformation-related Classification Tasks</i> Nikolaos Nikolaidis, Jakub Piskorski and Nicolas Stefanovitch.....	6992
<i>Extending AZee with Non-manual Gesture Rules for French Sign Language</i> Camille Challant and Michael Filhol.....	7007
<i>Extending the Discourse Analysis Tool Suite with Whiteboards for Visual Qualitative Analysis</i> Tim Fischer, Florian Schneider, Fynn Petersen-Frey, Anja Silvia Mollah Haque, Isabel Eiser, Gertraud Koch and Chris Biemann	7017
<i>Extracting Biomedical Entities from Noisy Audio Transcripts</i> Nima Ebadi, Kellen Morgan, Adrian Tan, Billy Linares, Sheri Osborn, Emma Majors, Jeremy Davis and Anthony Rios.....	7023
<i>Extracting Financial Events from Raw Texts via Matrix Chunking</i> Yusheng Huang, Ning Hu, Kunping Li, Nan Wang and Zhouhan Lin.....	7035
<i>Extracting Social Determinants of Health from Pediatric Patient Notes Using Large Language Models: Novel Corpus and Methods</i> Yujuan Fu, Giridhar Kaushik Ramachandran, Nicholas J. Dobbins, Namu Park, Michael Leu, Abby R. Rosenberg, Kevin Lybarger, Fei Xia, Özlem Uzuner and Meliha Yetisgen ...	7045
<i>Eye-Tracking Features Masking Transformer Attention in Question-Answering Tasks</i> Leran Zhang and Nora Hollenstein	7057
<i>Factorized Learning Assisted with Large Language Model for Gloss-free Sign Language Translation</i> Zhigang Chen, Benjia Zhou, Jun Li, Jun Wan, Zhen Lei, Ning Jiang, Quan Lu and Guoqing Zhao.....	7071
<i>FaGANet: An Evidence-Based Fact-Checking Model with Integrated Encoder Leveraging Contextual Information</i> Weiyao Luo, Junfeng Ran, Zailong Tian, Sujian Li and Zhifang Sui.....	7082
<i>FaiMA: Feature-aware In-context Learning for Multi-domain Aspect-based Sentiment Analysis</i> Songhua Yang, Xinke Jiang, Hanjie Zhao, Wenxuan Zeng, Hongde Liu and Yuxiang Jia	7089
<i>FAIRification of LeiLanD</i> Eric Sanders, Sara Petrollino, Gilles R. Scheifer, Henk van den Heuvel and Christopher Handy	7101

<i>FaAI: A Dataset for End-to-end Spoken Language Understanding in a Low-Resource Scenario</i> Andres Pineiro-Martin, Carmen Garcia-Mateo, Laura Docio-Fernandez, Maria del Carmen Lopez-Perez and Jose Gandarela-Rodriguez	7107
<i>Fast Adaptation via Prompted Data: An Efficient Cross-Domain Fine-tuning Method for Large Language Models</i> Yiming Zhang, Hantao Yang, Haobo Wang and Jake Zhao	7117
<i>FastSpell: The Langld Magic Spell</i> Marta Bañón, Gema Ramírez-Sánchez, Jaume Zaragoza-Bernabeu and Sergio Ortiz Rojas	7133
<i>FCDS: Fusing Constituency and Dependency Syntax into Document-Level Relation Extraction</i> Xudong Zhu, Zhao Kang and Bei Hui	7141
<i>Feature Structure Matching for Multi-source Sentiment Analysis with Efficient Adaptive Tuning</i> Rui Li, Cheng Liu, Yu Tong and Jiang Dazhi	7153
<i>Federated Document-Level Biomedical Relation Extraction with Localized Context Contrast</i> Yan Xiao, Yaochu Jin and Kuangrong Hao	7163
<i>Federated Foundation Models: Privacy-Preserving and Collaborative Learning for Large Models</i> Sixing Yu, Juan Pablo Munoz and Ali Jannesari	7174
<i>Few-Shot Learning for Cold-Start Recommendation</i> Mingming Li, Songlin Hu, Fuqing Zhu and Qiannan Zhu	7185
<i>Few-shot Link Prediction on Hyper-relational Facts</i> Jiyao Wei, Saiping Guan, Xiaolong Jin, Jiafeng Guo and Xueqi Cheng	7196
<i>Few-Shot Multimodal Named Entity Recognition Based on Mutlimodal Causal Intervention Graph</i> Feihong Lu, Xiaocui Yang, Qian Li, Qingyun Sun, Ke Jiang, Cheng Ji and Jianxin Li.	7208
<i>Few-shot Named Entity Recognition via Superposition Concept Discrimination</i> Jiawei Chen, Hongyu Lin, Xianpei Han, Yaojie Lu, Shanshan Jiang, Bin Dong and Le Sun	7220
<i>Few-Shot Relation Extraction with Hybrid Visual Evidence</i> Jiaying Gong and Hoda Eldardiry	7232
<i>Few-Shot Semantic Dependency Parsing via Graph Contrastive Learning</i> Bin Li, Yunlong Fan, Yikemaiti Sataer, Chuanqi Shi, Miao Gao and Zhiqiang Gao....	7248
<i>Few-shot Temporal Pruning Accelerates Diffusion Models for Text Generation</i> Bocheng Li, Zhujin Gao, Yongxin Zhu, Kun Yin, Haoyu Cao, Deqiang Jiang and Linli Xu	7259
<i>FFSTC: Fongbe to French Speech Translation Corpus</i> D. Fortuné Kponou, Fréjus A. A. Laleye and Eugène Cokou Ezin	7270
<i>FinCorpus-DE10k: A Corpus for the German Financial Domain</i> Serhii Hamotskyi, Nata Kozaeva and Christian Hänig	7277

<i>Finding Educationally Supportive Contexts for Vocabulary Learning with Attention-Based Models</i>	Sungjin Nam, Kevyn Collins-Thompson, David Jurgens and Xin Tong	7286
<i>Finding Spoken Identifications: Using GPT-4 Annotation for an Efficient and Fast Dataset Creation Pipeline</i>	Maliha Jahan, Helin Wang, Thomas Thebaud, Yinglun Sun, Giang Ha Le, Zsuzsanna Fagyal, Odette Scharenborg, Mark Hasegawa-Johnson, Laureano Moro Velazquez and Najim Dehak	7296
<i>Find-the-Common: A Benchmark for Explaining Visual Patterns from Images</i>	Yuting Shi, Naoya Inoue, Houjing Wei, Yufeng Zhao and Tao Jin	7307
<i>Fine-grained Classification of Circumstantial Meanings within the Prague Dependency Treebank Annotation Scheme</i>	Marie Mikulova	7314
<i>Fine-Grained Legal Argument-Pair Extraction via Coarse-Grained Pre-training</i>	Chaojun Xiao, Yutao Sun, Yuan Yao, Xu Han, Wenbin Zhang, Zhiyuan Liu and Maosong Sun	7324
<i>Fine-Tuning a Pre-Trained Wav2Vec2 Model for Automatic Speech Recognition- Experiments with De Zahrar Sproche</i>	Andrea Gulli, Francesco Costantini, Diego Sidraschi and Emanuela Li Destri	7336
<i>First Steps Towards the Integration of Resources on Historical Glossing Traditions in the History of Chinese: A Collection of Standardized Fǎnqiè Spellings from the Guǎngyùn</i>	Michele Pulini and Johann-Mattis List	7343
<i>Fisher Mask Nodes for Language Model Merging</i>	Thennal D K, Ganesh Nathan and Suchithra M S	7349
<i>FlattenQuant: Breaking through the Inference Compute-bound for Large Language Models with Per-tensor Quantization</i>	Yi Zhang, Fei Yang, Shuang Peng, Fangyu Wang and Aimin Pan	7356
<i>Flexible Lexicalization in Rule-based Text Realization</i>	Avril Gazeau and Francois Lareau	7366
<i>FLOR: On the Effectiveness of Language Adaptation</i>	Severino Da Dalt, Joan Llop, Irene Baucells, Marc Pamies, Yishi Xu, Aitor Gonzalez-Agirre and Marta Villegas	7377
<i>FoRC4CL: A Fine-grained Field of Research Classification and Annotated Dataset of NLP Articles</i>	Raia Abu Ahmad, Ekaterina Borisova and Georg Rehm	7389
<i>FORECAST2023: A Forecast and Reasoning Corpus of Argumentation Structures</i>	Kamila Górska, John Lawrence and Chris Reed	7395
<i>FoTo: Targeted Visual Topic Modeling for Focused Analysis of Short Texts</i>	Sanuj Kumar and Tuan Le	7406

<i>FRACAS: a FRENch Annotated Corpus of Attribution relations in newS</i> Ange Richard, Laura Cristina Alonzo Canul and François Portet.....	7417
<i>Frame2: A FrameNet-based Multimodal Dataset for Tackling Text-image Interactions in Video</i> Frederico Belcavello, Tiago Timponi Torrent, Ely E. Matos, Adriana S. Pagano, Maucha Gamonal, Natalia Sigiliano, Livia Vicente Dutra, Helen de Andrade Abreu, Mairon Samagaio, Mariane Carvalho, Franciany Campos, Gabrielly Azalim, Bruna Mazzei, Mateus Fonseca de Oliveira, Ana Carolina Loçasso Luz, Livia Pádua Ruiz, Júlia Bellei, Amanda Pestana, Josiane Costa, Iasmin Rabelo, Anna Beatriz Silva, Raquel Roza, Mariana Souza and Igor Oliveira.....	7429
<i>Framed Multi30K: A Frame-Based Multimodal-Multilingual Dataset</i> Marcelo Viridiano, Arthur Lorenzi, Tiago Timponi Torrent, Ely E. Matos, Adriana S. Pagano, Natália Sathler Sigiliano, Maucha Gamonal, Helen de Andrade Abreu, Livia Vicente Dutra, Mairon Samagaio, Mariane Carvalho, Franciany Campos, Gabrielly Azalim, Bruna Mazzei, Mateus Fonseca de Oliveira, Ana Carolina Luz, Livia Padua Ruiz, Júlia Bellei, Amanda Pestana, Josiane Costa, Iasmin Rabelo, Anna Beatriz Silva, Raquel Roza, Mariana Souza Mota, Igor Oliveira and Márcio Henrique Pelegrino de Freitas	7438
<i>FRASIMED: A Clinical French Annotated Resource Produced through Crosslingual BERT-Based Annotation Projection</i> Jamil Zaghir, Mina Bjelogrljic, Jean-Philippe Goldman, Soukaïna Aananou, Christophe Gaudet-Blavignac and Christian Lovis	7450
<i>FReND: A French Resource of Negation Data</i> Hafida Le Cloirec - Ait Yahya, Olga Seminck and Pascal Amsili	7461
<i>From Graph to Word Bag: Introducing Domain Knowledge to Confusing Charge Prediction</i> Ang Li, Qiangchao Chen, Yiquan Wu, Xiang Zhou, Kun Kuang, Fei Wu and Ming Cai.....	7469
<i>From Laughter to Inequality: Annotated Dataset for Misogyny Detection in Tamil and Malayalam Memes</i> Rahul Ponnusamy, Kathiravan Pannerselvam, Saranya R, Prasanna Kumar Kumaresan, Sajeetha Thavareesan, Bhuvaneswari S, Anshid K.A, Susminu S Kumar, Paul Buitelaar and Bharathi Raja Chakravarthi.....	7480
<i>From Linguistic Linked Data to Big Data</i> Dimitar Trajanov, Elena Apostol, Radovan Garabik, Katerina Gkirtzou, Dagmar Gromann, Chaya Liebeskind, Cosimo Palma, Michael Rosner, Alexia Sampri, Gilles Sérasset, Blerina Spahiu, Ciprian-Octavian Truică and Giedre Valunaite Oleskeviciene	7489
<i>From News to Summaries: Building a Hungarian Corpus for Extractive and Abstractive Summarization</i> Botond Barta, Dorina Lakatos, Attila Nagy, Milán Konor Nyist and Judit Ács.....	7503
<i>From Technology to Market. Bilingual Corpus on the Evaluation of Technology Opportunity Discovery</i> Amir Hazem, Kazuyuki Motohashi and Chen Zhu	7510
<i>From Text to Historical Ecological Knowledge: The Construction and Application of the Shan Jing Knowledge Base</i> Ke Liang, Chu-Ren Huang and Xin-Lan Jiang.....	7521

<i>From Text to Source: Results in Detecting Large Language Model-Generated Content</i> Wissam Antoun, Benoît Sagot and Djamé Seddah	7531
<i>FUSE - Frustration and Surprise Expressions: A Subtle Emotional Multimodal Language Corpus</i> Rajesh Titung and Cecilia Ovesdotter Alm	7544
<i>Fusion-in-T5: Unifying Variant Signals for Simple and Effective Document Ranking with Attention Fusion</i> Shi Yu, Chenghao Fan, Chenyan Xiong, David Jin, Zhiyuan Liu and Zhenghao Liu ..	7556
<i>GAATME: A Genetic Algorithm for Adversarial Translation Metrics Evaluation</i> Josef Jon and Ondřej Bojar	7562
<i>GCNet: Global-and-Context Collaborative Learning for Aspect-Based Sentiment Analysis</i> Ting Zhou, Ying Shen and Yinghui Li	7570
<i>GECSum: Generative Evaluation-Driven Sequence Level Contrastive Learning for Abstractive Summarization</i> Jiawen Xie, Shaoting Zhang and Xiaofan Zhang	7581
<i>Gendered Grammar or Ingrained Bias? Exploring Gender Bias in Icelandic Language Models</i> Steinunn Rut Friðriksdóttir and Hafsteinn Einarsson	7596
<i>Generating Clarification Questions for Disambiguating Contracts</i> Anmol Singhal, Chirag Jain, Preethu Rose Anish, Arkajyoti Chakraborty and Smita Ghaisas	7611
<i>Generating Contextual Images for Long-Form Text</i> Avijit Mitra, Nalin Gupta, Chetan Naik, Abhinav Sethy, Kinsey Bice and Zeynab Raeesy	7623
<i>Generating Hard-Negative Out-of-Scope Data with ChatGPT for Intent Classification</i> Zhijian Li, Stefan Larson and Kevin Leach	7634
<i>Generating Multiple-choice Questions for Medical Question Answering with Distractors and Cue-masking</i> Damien Sileo, Kanimozhi Uma and Marie-Francine Moens	7647
<i>Generative Multimodal Entity Linking</i> senbao shi, Zhenran Xu, Baotian Hu and Min Zhang	7654
<i>GENTRAC: A Tool for Tracing Trauma in Genocide and Mass Atrocity Court Transcripts</i> Miriam Schirmer, Christian Brechenmacher and Juergen Pfeffer	7666
<i>Geographically-Informed Language Identification</i> Jonathan Dunn and Lane Edwards-Brown	7672
<i>GerDISDETECT: A German Multilabel Dataset for Disinformation Detection</i> Mina Schütz, Daniela PISOIU, Daria Liakhovets, Alexander Schindler and Melanie Siegel	7683
<i>German Also Hallucinates! Inconsistency Detection in News Summaries with the Absinth Dataset</i> Laura Mascarell, Ribin Chalumattu and Annette Rios	7696

<i>German Parliamentary Corpus (GerParCor) Reloaded</i> Giuseppe Abrami, Mevlüt Bağcı and Alexander Mehler	7707
<i>German SRL: Corpus Construction and Model Training</i> Maxim Konca, Andy Luecking and Alexander Mehler	7717
<i>GERMS-AT: A Sexism/Misogyny Dataset of Forum Comments from an Austrian Online Newspaper</i> Brigitte Krenn, Johann Petrak, Marina Kubina and Christian Burger	7728
<i>GIL-GALaD: Gender Inclusive Language - German Auto-Assembled Large Database</i> Anna-Katharina Dick, Matthias Drews, Valentin Pickard and Victoria Pierz	7740
<i>GLAMR: Augmenting AMR with GL-VerbNet Event Structure</i> Jingxuan Tu, Timothy Obiso, Bingyang Ye, Kyeongmin Rim, Keer Xu, Liulu Yue, Susan Windisch Brown, Martha Palmer and James Pustejovsky	7746
<i>Global and Local Hierarchical Prompt Tuning Framework for Multi-level Implicit Discourse Relation Recognition</i> Lei Zeng, Ruifang He, Haowen Sun, Jing Xu, Chang Liu and Bo Wang	7760
<i>GlotScript: A Resource and Tool for Low Resource Writing System Identification</i> Amir Hossein Kargaran, François Yvon and Hinrich Schütze	7774
<i>GMEG-EXP: A Dataset of Human- and LLM-Generated Explanations of Grammatical and Fluency Edits</i> S. Magalí López Cortez, Mark Josef Norris and Steve Duman	7785
<i>GOLEM: GOld Standard for Learning and Evaluation of Motifs</i> W. Victor Yarlott, Anurag Acharya, Diego Castro Estrada, Diana Gomez and Mark Finlayson	7801
<i>Good or Bad News? Exploring GPT-4 for Sentiment Analysis for Faroese on a Public News Corpora</i> Iben Nyholm Debess, Annika Simonsen and Hafsteinn Einarsson	7814
<i>Gos 2: A New Reference Corpus of Spoken Slovenian</i> Darinka Verdonik, Kaja Dobrovoljc, Tomaž Erjavec and Nikola Ljubešić	7825
<i>GPT-3.5 for Grammatical Error Correction</i> Anisia Katinskaia and Roman Yangarber	7831
<i>GPTEval: A Survey on Assessments of ChatGPT and GPT-4</i> Rui Mao, Guanyi Chen, Xulang Zhang, Frank Guerin and Erik Cambria	7844
<i>GPT-HateCheck: Can LLMs Write Better Functional Tests for Hate Speech Detection?</i> Yiping Jin, Leo Wanner and Alexander Shvets	7867
<i>GPT-SW3: An Autoregressive Language Model for the Scandinavian Languages</i> Ariel Ekgren, Amaru Cuba Gyllensten, Felix Stollenwerk, Joey Öhman, Tim Isbister, Evangelia Gogoulou, Fredrik Carlsson, Judit Casademont and Magnus Sahlgren	7886

<i>Gradient Consistency-based Parameter Allocation for Multilingual Neural Machine Translation</i> Wenshuai Huo, Xiaocheng Feng, Yichong Huang, Chengpeng Fu, Hui Wang and Bing Qin	7901
<i>Gramble: A Tabular Programming Language for Collaborative Linguistic Modeling</i> Patrick Littell, Darlene Stewart, Fineen Davis, Aidan Pine and Roland Kuhn	7913
<i>Grammatical Error Correction for Code-Switched Sentences by Learners of English</i> Kelvin Wey Han Chan, Christopher Bryant, Li Nguyen, Andrew Caines and Zheng Yuan	7926
<i>Granular Change Accuracy: A More Accurate Performance Metric for Dialogue State Tracking</i> Taha Aksu and Nancy Chen	7939
<i>GreekBART: The First Pretrained Greek Sequence-to-Sequence Model</i> Iakovos Evdaimon, Hadi Abdine, Christos Xypolopoulos, Stamatis Outsios, Michalis Vazirgiannis and Giorgos Stamou	7949
<i>GRIT: A Dataset of Group Reference Recognition in Italian</i> Sergio E. Zannotto, Qi Yu, Miriam Butt and Diego Frassinelli	7963
<i>Grounded Multimodal Procedural Entity Recognition for Procedural Documents: A New Dataset and Baseline</i> Haopeng Ren, Yushi Zeng, Yi Cai, Zhenqi Ye, Li Yuan and Pinli Zhu	7971
<i>Guided Distant Supervision for Multilingual Relation Extraction Data: Adapting to a New Language</i> Alistair Plum, Tharindu Ranasinghe and Christoph Purschke	7982
<i>HAE-RAE Bench: Evaluation of Korean Knowledge in Language Models</i> Guijin Son, Hanwool Lee, suwan kim, Huiseo Kim, Jae cheol Lee, Je Won Yeom, Jihyu Jung, Jung woo Kim and Songseong Kim	7993
<i>Halwasa: Quantify and Analyze Hallucinations in Large Language Models: Arabic as a Case Study</i> Hamdy Mubarak, Hend Al-Khalifa and Khaloud Suliman Alkhalefah	8008
<i>HarmPot: An Annotation Framework for Evaluating Offline Harm Potential of Social Media Text</i> Ritesh Kumar, Ojaswee Bhalla, Madhu Vanthi, Shehlat Maknoon Wani and Siddharth Singh	8016
<i>Harnessing the Power of Large Language Model for Uncertainty Aware Graph Processing</i> Zhenyu Qian, Yiming Qian, Yuting Song, Fei Gao, Hai Jin, Chen Yu and Xia Xie	8035
<i>Has It All Been Solved? Open NLP Research Questions Not Solved by Large Language Models</i> Oana Ignat, Zhijing Jin, Artem Abzaliev, Laura Biester, Santiago Castro, Naihao Deng, Xinyi Gao, Aylin Ece Gunal, Jacky He, Ashkan Kazemi, Muhammad Khalifa, Namho Koh, Andrew Lee, Siyang Liu, Do June Min, Shinka Mori, Joan C. Nwatu, Veronica Perez-Rosas, Siqi Shen, Zekun Wang, Winston Wu and Rada Mihalcea	8050
<i>HealthFC: Verifying Health Claims with Evidence-Based Medical Fact-Checking</i> Juraj Vladika, Phillip Schneider and Florian Matthes	8095

<i>Hierarchical Graph Convolutional Network Approach for Detecting Low-Quality Documents</i> Jaeyoung Lee, Joonwon Jang and Misuk Kim.....	8108
<i>Hierarchical Selection of Important Context for Generative Event Causality Identification with Optimal Transports</i> Hieu Man, Chien Van Nguyen, Nghia Trung Ngo, Linh Ngo, Franck Dernoncourt and Thien Huu Nguyen	8122
<i>Hierarchical Topic Modeling via Contrastive Learning and Hyperbolic Embedding</i> Zhicheng Lin, HeGang Chen, Yuyin Lu, Yanghui Rao, Hao Xu and Hanjiang Lai.....	8133
<i>High-order Joint Constituency and Dependency Parsing</i> Yanggan Gu, Yang Hou, Zhefeng Wang, Xinyu Duan and Zhenghua Li.....	8144
<i>High-Order Semantic Alignment for Unsupervised Fine-Grained Image-Text Retrieval</i> Rui Gao, Miaomiao Cheng, Xu Han and Wei Song.....	8155
<i>HoLM: Analyzing the Linguistic Unexpectedness in Homeric Poetry</i> John Pavlopoulos, Ryan Sandell, Maria Konstantinidou and Chiara Bozzone	8166
<i>How Diplomats Dispute: The UN Security Council Conflict Corpus</i> Karolina Zaczynska, Peter Bourgonje and Manfred Stede	8173
<i>How Do Hyenas Deal with Human Speech? Speech Recognition and Translation with ConfHyena</i> Marco Gaido, Sara Papi, Matteo Negri and Luisa Bentivogli.....	8184
<i>How Far Is Too Far? Studying the Effects of Domain Discrepancy on Masked Language Models</i> Subhradeep Kayal, Alexander Rakhlin, Ali Dashti and Serguei Stepaniants.....	8192
<i>How Gender Interacts with Political Values: A Case Study on Czech BERT Models</i> Adnan Al Ali and Jindřich Libovický	8200
<i>How Good Are LLMs at Out-of-Distribution Detection?</i> Bo Liu, Li-Ming Zhan, Zexin Lu, Yujie Feng, Lei Xue and Xiao-Ming Wu	8211
<i>How Important Is Tokenization in French Medical Masked Language Models?</i> Yanis Labrak, Adrien Bazoge, Béatrice Daille, Mickael Rouvier and Richard Dufour .	8223
<i>How Large Language Models Encode Context Knowledge? A Layer-Wise Probing Study</i> Tianjie Ju, Weiwei Sun, Wei Du, Xinwei Yuan, Zhaochun Ren and Gongshen Liu....	8235
<i>How Much Do Robots Understand Rudeness? Challenges in Human-Robot Interaction</i> Michael Andrew Orme, Yanchao Yu and Zhiyuan Tan	8247
<i>How Robust Are the QA Models for Hybrid Scientific Tabular Data? A Study Using Customized Dataset</i> Akash Ghosh, Venkata Sahith Bathini, Niloy Ganguly, Pawan Goyal and Mayank Singh	8258
<i>How Speculative Can Speculative Decoding Be?</i> Zhuorui Liu, Chen Zhang and Dawei Song.....	8265
<i>How Susceptible Are LLMs to Logical Fallacies?</i> Amirreza Payandeh, Dan Pluth, Jordan Hosier, Xuesu Xiao and Vijay K. Gurbani....	8276

<i>How to Do Politics with Words: Investigating Speech Acts in Parliamentary Debates</i> Ines Reinig, Ines Rehbein and Simone Paolo Ponzetto.....	8287
<i>How to Encode Domain Information in Relation Classification</i> Elisa Bassignana, Viggo Unmack Gascou, Frida Nøhr Laustsen, Gustav Kristensen, Marie Haahr Petersen, Rob van der Goot and Barbara Plank.....	8301
<i>How to Solve Few-Shot Abusive Content Detection Using the Data We Actually Have</i> Viktor Hangya and Alexander Fraser.....	8307
<i>How to Understand "Support"? An Implicit-enhanced Causal Inference Approach for Weakly-supervised Phrase Grounding</i> Jiamin Luo, Jianing Zhao, Jingjing Wang and Guodong Zhou.....	8323
<i>How Well Can BERT Learn the Grammar of an Agglutinative and Flexible-Order Language? The Case of Basque.</i> Gorka Urbizu, Maitze Zulaika, Xabier Saralegi and Ander Corral.....	8334
<i>HS-GC: Holistic Semantic Embedding and Global Contrast for Effective Text Clustering</i> Chen Yang, Bin Cao and Jing Fan.....	8349
<i>HuLU: Hungarian Language Understanding Benchmark Kit</i> Noémi Ligeti-Nagy, Gergő Ferenczi, Enikő Héja, László János Laki, Noémi Vadász, Zijian Győző Yang and Tamás Váradi.....	8360
<i>Human and System Perspectives on the Expression of Irony: An Analysis of Likelihood Labels and Rationales</i> Aaron Maladry, Alessandra Teresa Cignarella, Els Lefever, Cynthia van Hee and Veronique Hoste.....	8372
<i>HumanEval-XL: A Multilingual Code Generation Benchmark for Cross-lingual Natural Language Generalization</i> Qiwei Peng, Yekun Chai and Xuhong Li.....	8383
<i>Human in the Loop: How to Effectively Create Coherent Topics by Manually Labeling Only a Few Documents per Class</i> Anton F. Thielmann, Christoph Weisser and Benjamin Säfken.....	8395
<i>Humanistic Buddhism Corpus: A Challenging Domain-Specific Dataset of English Translations for Classical and Modern Chinese</i> Youheng W. Wong, Natalie Parde and Erdem Koyuncu.....	8406
<i>Humanitarian Corpora for English, French and Spanish</i> Loryn Isaacs, Santiago Chambó and Pilar León-Araúz.....	8418
<i>Humanizing Machine-Generated Content: Evading AI-Text Detection through Adversarial Attack</i> Ying Zhou, Ben He and Le Sun.....	8427
<i>Humans Need Context, What about Machines? Investigating Conversational Context in Abusive Language Detection</i> Tom Bourgeade, Zongmin Li, Farah Benamara, Véronique Moriceau, Jian Su and Aixin Sun.....	8438

<i>Human vs. Machine Perceptions on Immigration Stereotypes</i> Wolfgang S. Schmeisser-Nieto, Pol Pastells, Simona Frenda and Mariona Taule	8453
<i>Hybrid of Spans and Table-Filling for Aspect-Level Sentiment Triplet Extraction</i> Minghua Nuo and Chaofan Guo	8464
<i>Hyperbolic Graph Neural Network for Temporal Knowledge Graph Completion</i> Yancong Li, Xiaoming Zhang, Ying Cui and Shuai Ma	8474
<i>Hyperbolic Representations for Prompt Learning</i> Nan Chen, Xiangdong Su and Feilong Bao	8487
<i>Hypergraph-Based Session Modeling: A Multi-Collaborative Self-Supervised Approach for Enhanced Recommender Systems</i> Xiangping Zheng, Bo Wu, Alex X. Zhang and Wei Li	8493
<i>HyperMR: Hyperbolic Hypergraph Multi-hop Reasoning for Knowledge-based Visual Question Answering</i> Bin Wang, Fuyong Xu, Peiyu Liu and Zhenfang Zhu	8505
<i>HYPERTTS: Parameter Efficient Adaptation in Text to Speech Using Hypernetworks</i> Yingting Li, Rishabh Bhardwaj, Ambuj Mehrish, Bo Cheng and Soujanya Poria	8516
<i>HYRR: Hybrid Infused Reranking for Passage Retrieval</i> Jing Lu, Keith Hall, Ji Ma and Jianmo Ni	8528
<i>IAD: In-Context Learning Ability Decoupler of Large Language Models in Meta-Training</i> Yuhan Liu, Xiuying Chen, Gao Xing, Ji Zhang and Rui Yan	8535
<i>IDC: Boost Text-to-image Retrieval via Indirect and Direct Connections</i> Guowei Ge, Kuangrong Hao and Linguang Hao	8546
<i>IDEATE: Detecting AI-Generated Text Using Internal and External Factual Structures</i> Quan Wang, Licheng Zhang, Zikang Guo and Zhendong Mao	8556
<i>IDEM: The IDioms with EMotions Dataset for Emotion Recognition</i> Alexander Prochnow, Johannes E. Bendler, Caroline Lange, Foivos Ioannis Tzavellos, Bas Marco Göritzer, Marijn ten Thij and Riza Batista-Navarro	8569
<i>Identifying and Aligning Medical Claims Made on Social Media with Medical Evidence</i> Anthony James Hughes and Xingyi Song	8580
<i>Identifying Fine-grained Depression Signs in Social Media Posts</i> Augusto R. Mendes and Helena Caseli	8594
<i>Identifying Source Language Expressions for Pre-editing in Machine Translation</i> Norizo Sakaguchi, Yugo Murawaki, Chenhui Chu and Sadao Kurohashi	8605
<i>Ideological Knowledge Representation: Framing Climate Change in EcoLexicon</i> Arianne Reimerink, Melania Cabezas-García, Pilar León-Araúz and Pamela Faber	8617
<i>ILCiteR: Evidence-grounded Interpretable Local Citation Recommendation</i> Sayar Ghosh Roy and Jiawei Han	8627

<i>ILLUMINER: Instruction-tuned Large Language Models as Few-shot Intent Classifier and Slot Filler</i>	
Paramita Mirza, Viju Sudhi, Soumya Ranjan Sahoo and Sinchana Ramakanth Bhat .	8639
<i>Image Matters: A New Dataset and Empirical Study for Multimodal Hyperbole Detection</i>	
Huixuan Zhang and Xiaojun Wan	8652
<i>Impact of Task Adapting on Transformer Models for Targeted Sentiment Analysis in Croatian Headlines</i>	
Sofia Lee and Jelke Bloem	8662
<i>Impoverished Language Technology: The Lack of (Social) Class in NLP</i>	
Amanda Cercas Curry, Zeerak Talat and Dirk Hovy	8675
<i>Improved Neural Protoform Reconstruction via Reflex Prediction</i>	
Liang Lu, Jingzhi Wang and David R. Mortensen.....	8683
<i>Improved Out-of-Scope Intent Classification with Dual Encoding and Threshold-based Re-Classification</i>	
Hossam Zawbaa, Wael Rashwan, Sourav Dutta and Haytham Assem.....	8708
<i>Improving Bengali and Hindi Large Language Models</i>	
Arif Shahriar and Denilson Barbosa.....	8719
<i>Improving Chinese Named Entity Recognition with Multi-grained Words and Part-of-Speech Tags via Joint Modeling</i>	
Chenhui Dou, Chen Gong, Zhenghua Li, Zhefeng Wang, baoxing Huai and Min Zhang	8732
<i>Improving Content Recommendation: Knowledge Graph-Based Semantic Contrastive Learning for Diversity and Cold-Start Users</i>	
Yejin Kim, Scott Rome, Kevin Foley, Mayur Nankani, Rimon Melamed, Javier Morales, Abhay K. Yadav, Maria Peifer, Sardar Hamidian and H. Howie Huang	8743
<i>Improving Continual Few-shot Relation Extraction through Relational Knowledge Distillation and Prototype Augmentation</i>	
Zhiheng Zhang, Daojian Zeng and Xue Bai	8756
<i>Improving Copy-oriented Text Generation via EDU Copy Mechanism</i>	
Tianxiang Wu, Han Chen, Luozheng Qin, Ziqiang Cao and Chunhui Ai.....	8768
<i>Improving Cross-lingual Transfer with Contrastive Negative Learning and Self-training</i>	
Guanlin Li, Xuechen Zhao, Amir Jafari, Wenhao Shao, Reza Farahbakhsh and Noel Crespi	8781
<i>Improving Factual Consistency in Abstractive Summarization with Sentence Structure Pruning</i>	
Dingxin Hu, Xuanyu Zhang, Xingyue Zhang, Yiyang Li, Dongsheng Chen, Marina Litvak, Natalia Vanetik, Qing Yang, Dongliang Xu, Yanquan Zhou, Lei Li, Yuze Li and Yingqi Zhu	8792
<i>Improving Faithfulness of Large Language Models in Summarization via Sliding Generation and Self-Consistency</i>	
Taiji Li, Zhi Li and Yin Zhang	8804
<i>Improving Grammatical Error Correction by Correction Acceptability Discrimination</i>	
Bin Cao, Kai Jiang, Fayu Pan, Chenlei Bao and Jing Fan	8818

<i>Improving Implicit Discourse Relation Recognition with Semantics Confrontation</i> Mingyang Cai, Zhen Yang and Ping Jian	8828
<i>Improving Language Model Reasoning with Self-motivated Learning</i> Yunlong Feng, Yang Xu, Libo Qin, Yasheng Wang and Wanxiang Che	8840
<i>Improving Low-Resource Keyphrase Generation through Unsupervised Title Phrase Generation</i> Byungha Kang and Youhyun Shin	8853
<i>Improving Multi-view Document Clustering: Leveraging Multi-structure Processor and Hybrid Ensemble Clustering Module</i> Ruina Bai and Qi Bai	8866
<i>Improving Personalized Sentiment Representation with Knowledge-enhanced and Parameter-efficient Layer Normalization</i> You Zhang, Jin Wang, Liang-Chih Yu, Dan Xu and Xuejie Zhang	8877
<i>Improving Recall of Large Language Models: A Model Collaboration Approach for Relational Triple Extraction</i> Zepeng Ding, Wenhao Huang, Jiaqing Liang, Yanghua Xiao and Deqing Yang	8890
<i>Improving Robustness of GNN-based Anomaly Detection by Graph Adversarial Training</i> Xiangping Zheng, Bo Wu, Alex X. Zhang and Wei Li	8902
<i>Improving Role-Oriented Dialogue Summarization with Interaction-Aware Contrastive Learning</i> Weihong Guan, Shi Feng, Daling Wang, Faliang Huang, Yifei Zhang and Yuan Cui ..	8913
<i>Improving Text Readability through Segmentation into Rhemes</i> Antoine Jamelot, Solen Quiniou and Sophie Hamon	8925
<i>Improving the Robustness of Large Language Models via Consistency Alignment</i> Yukun Zhao, Lingyong Yan, Weiwei Sun, Guoliang Xing, Shuaiqiang Wang, Chong Meng, Zhicong Cheng, Zhaochun Ren and Dawei Yin	8931
<i>Improving Unsupervised Neural Machine Translation via Training Data Self-Correction</i> Jinliang Lu and Jiajun Zhang	8942
<i>Improving Vietnamese-English Medical Machine Translation</i> Nhu Vo, Dat Quoc Nguyen, Dung D. Le, Massimo Piccardi and Wray Buntine	8955
<i>InaGVAD : A Challenging French TV and Radio Corpus Annotated for Speech Activity Detection and Speaker Gender Segmentation</i> David Doukhan, Christine Maertens, William Le Personnic, Ludovic Speroni and Reda Dehak	8963
<i>In-Context Example Retrieval from Multi-Perspectives for Few-Shot Aspect-Based Sentiment Analysis</i> Qianlong Wang, Hongling Xu, Keyang Ding, Bin Liang and Ruifeng Xu	8975
<i>Incorporating Lexical and Syntactic Knowledge for Unsupervised Cross-Lingual Transfer</i> Jianyu Zheng, Fengfei Fan and Jianquan Li	8986

<i>Incorporating Word-level Phonemic Decoding into Readability Assessment</i> Christine Pinney, Casey Kennington, Maria Soledad Pera, Katherine Landau Wright and Jerry Alan Fails	8998
<i>IndicFinNLP: Financial Natural Language Processing for Indian Languages</i> Sohom Ghosh, Arnab Maji, Aswartha Narayana and Sudip Kumar Naskar	9010
<i>Indic-TEDST: Datasets and Baselines for Low-Resource Speech to Text Translation</i> Nivedita Sethiya, Saanvi Nair and Chandresh Maurya.....	9019
<i>IndirectQA: Understanding Indirect Answers to Implicit Polar Questions in French and Spanish</i> Christin Müller and Barbara Plank	9025
<i>Inductive Knowledge Graph Completion with GNNs and Rules: An Analysis</i> Akash Anil, Victor Gutierrez-Basulto, Yazmin Ibanez-Garcia and Steven Schockaert.	9036
<i>InferBR: A Natural Language Inference Dataset in Portuguese</i> Luciana Bencke, Francielle Vasconcellos Pereira, Moniele Kunrath Santos and Viviane Moreira	9050
<i>InfFeed: Influence Functions as a Feedback to Improve the Performance of Subjective Tasks</i> Somnath Banerjee, Maulindu Sarkar, Punyajoy Saha, Binny Mathew and Animesh Mukherjee.....	9061
<i>InfoEnh: Towards Multimodal Sentiment Analysis via Information Bottleneck Filter and Optimal Transport Alignment</i> Yifeng Xie, Zhihong Zhu, Xuan Lu, Zhiqi Huang and Haoran Xiong.....	9073
<i>Information Extraction with Differentiable Beam Search on Graph RNNs</i> Niama El Khbir, Nadi Tomeh and Thierry Charnois.....	9084
<i>INMT-Lite: Accelerating Low-Resource Language Data Collection via Offline Interactive Neural Machine Translation</i> Harshita Diddee, Anurag Shukla, Tanuja Ganu, Vivek Seshadri, Sandipan Dandapat, Monojit Choudhury and Kalika Bali	9097
<i>Integrating Headedness Information into an Auto-generated Multilingual CCGbank for Improved Semantic Interpretation</i> Tu-Anh Tran and Yusuke Miyao.....	9110
<i>Integrating Representation Subspace Mapping with Unimodal Auxiliary Loss for Attention-based Multimodal Emotion Recognition</i> Xulong Du, Xingnan Zhang, Dandan Wang, Yingying Xu, Zhiyuan Wu, Shiqing Zhang, Xiaoming Zhao, Jun Yu and Liangliang Lou.....	9120
<i>Intent-Aware and Hate-Mitigating Counterspeech Generation via Dual-Discriminator Guided LLMs</i> Haiyang Wang, Zhiliang Tian, Xin Song, Yue Zhang, Yuchen Pan, Hongkui Tu, Minlie Huang and Bin Zhou	9131
<i>Intention and Face in Dialog</i> Adil Soubki and Owen Rambow	9143

<i>InteRead: An Eye Tracking Dataset of Interrupted Reading</i> Francesca Zermiani, Prajit Dhar, Ekta Sood, Fabian Kögel, Andreas Bulling and Maria Wirzberger	9154
<i>Interpretable Assessment of Speech Intelligibility Using Deep Learning: A Case Study on Speech Disorders Due to Head and Neck Cancers</i> Sondes Abderrazek, Corinne Fredouille, Alain Ghio, Muriel Lalain, Christine Meunier, Mathieu Balaguer and Virginie Woisard	9170
<i>Interpretable Short Video Rumor Detection Based on Modality Tampering</i> Kaixuan Wu, Yanghao Lin, Donglin Cao and Dazhen Lin	9180
<i>Interpreting Themes from Educational Stories</i> Yigeng Zhang, Fabio Gonzalez and Thamar Solorio	9190
<i>Intrinsic Subgraph Generation for Interpretable Graph Based Visual Question Answering</i> Pascal Tilli and Ngoc Thang Vu	9204
<i>Introducing a Parsed Corpus of Historical High German</i> Christopher D. Sapp, Elliott Evans, Rex Sprouse and Daniel Dakota	9224
<i>Introducing CQuAE : A New French Contextualised Question-Answering Corpus for the Education Domain</i> Thomas Gerald, Anne Vilnat, Sofiane Ettayeb, Louis Tamames and Patrick Paroubek	9234
<i>Investigating the Robustness of Modelling Decisions for Few-Shot Cross-Topic Stance Detection: A Preregistered Study</i> Myrthe Reuver, Suzan Verberne and Antske Fokkens	9245
<i>IR2: Information Regularization for Information Retrieval</i> Jianyou Wang, Kaicheng Wang, Xiaoyue Wang, Weili Cao, Ramamohan Paturi and Leon Bergen	9261
<i>I Remember You!: SUI Corpus for Remembering and Utilizing Users' Information in Chat-oriented Dialogue Systems</i> Yuiko Tsunomori and Ryuichiro Higashinaka	9285
<i>ÌròyìnSpeech: A Multi-purpose Yorùbá Speech Corpus</i> Tolulope Ogunremi, Kola Tubosun, Anuluwapo Aremu, Iro Orife and David Ifeoluwa Adelani	9296
<i>Is Crowdsourcing Breaking Your Bank? Cost-Effective Fine-Tuning of Pre-trained Language Models with Proximal Policy Optimization</i> Shuo Yang and Gjergji Kasneci	9304
<i>Is Gender Reference Gender-specific? Studies in a Polar Domain</i> Manfred Klenner and Dylan Massey	9315
<i>Is It Possible to Modify Text to a Target Readability Level? An Initial Investigation Using Zero-Shot Large Language Models</i> Asma Farajidizaji, Vatsal Raina and Mark Gales	9325

<i>Is LLM a Reliable Reviewer? A Comprehensive Evaluation of LLM on Automatic Paper Reviewing Tasks</i>	
Ruiyang Zhou, Lu Chen and Kai Yu	9340
<i>Is Modularity Transferable? A Case Study through the Lens of Knowledge Distillation</i>	
Mateusz Klimaszewski, Piotr Andruszkiewicz and Alexandra Birch	9352
<i>ISO 24617-12: A New Standard for Semantic Annotation</i>	
Harry Bunt	9361
<i>IsraParlTweet: The Israeli Parliamentary and Twitter Resource</i>	
Guy Mor-Lan, Effi Levi, Tamir Sheaffer and Shaul R. Shenhav	9372
<i>Is Spoken Hungarian Low-resource?: A Quantitative Survey of Hungarian Speech Data Sets</i>	
Peter Mihajlik, Katalin Mády, Anna Kohári, Fruzsina Sára Fruzsina, Gábor Kiss, Tekla Etelka Gráczki and A. Seza Dođruöz	9382
<i>Is Summary Useful or Not? An Extrinsic Human Evaluation of Text Summaries on Downstream Tasks</i>	
Xiao Pu, Mingqi Gao and Xiaojun Wan	9389
<i>IT2ACL Learning Easy-to-Hard Instructions via 2-Phase Automated Curriculum Learning for Large Language Models</i>	
Yufei Huang and Deyi Xiong	9405
<i>IT5: Text-to-text Pretraining for Italian Language Understanding and Generation</i>	
Gabriele Sarti and Malvina Nissim	9422
<i>Italian Word Embeddings for the Medical Domain</i>	
Franco Alberto Cardillo and Franca Debole	9434
<i>It's Not under the Lamppost: Expanding the Reach of Conversational AI</i>	
Christy Doran and Deborah A. Dahl	9441
<i>JaParaPat: A Large-Scale Japanese-English Parallel Patent Application Corpus</i>	
Masaaki Nagata, Makoto Morishita, Katsuki Chousa and Norihito Yasuda	9452
<i>Jargon: A Suite of Language Models and Evaluation Tasks for French Specialized Domains</i>	
Vincent Segonne, Aidan Mannion, Laura Cristina Alonzo Canul, Alexandre Daniel Audibert, Xingyu Liu, Cécile Macaire, Adrien Pupier, Yongxin Zhou, Mathilde Aguiar, Felix E. Herron, Magali Norré, Massih R Amini, Pierrette Bouillon, Iris Eshkol-Taravella, Emmanuelle Esperança-Rodier, Thomas François, Lorraine Goeuriot, Jérôme Goulian, Mathieu Lafourcade, Benjamin Lecouteux, François Portet, Fabien Ringeval, Vincent Vandeghinste, Maximin Coavoux, Marco Dinarelli and Didier Schwab	9463
<i>JCoLA: Japanese Corpus of Linguistic Acceptability</i>	
Taiga Someya, Yushi Sugimoto and Yohei Oseki	9477
<i>J-CRe3: A Japanese Conversation Dataset for Real-world Reference Resolution</i>	
Nobuhiro Ueda, Hideko Habe, Akishige Yuguchi, Seiya Kawano, Yasutomo Kawanishi, Sadao Kurohashi and Koichiro Yoshino	9489
<i>JDocQA: Japanese Document Question Answering Dataset for Generative Language Models</i>	
Eri Onami, Shuhei Kurita, Taiki Miyanishi and Taro Watanabe	9503

<i>JEMHopQA: Dataset for Japanese Explainable Multi-Hop Question Answering</i> Ai Ishii, Naoya Inoue, Hisami Suzuki and Satoshi Sekine	9515
<i>JFLD: A Japanese Benchmark for Deductive Reasoning Based on Formal Logic</i> Terufumi Morishita, Atsuki Yamaguchi, Gaku Morio, Hikaru Tomonari, Osamu Imaichi and Yasuhiro Sogawa	9526
<i>JLBert: Japanese Light BERT for Cross-Domain Short Text Classification</i> Chandrai Kayal, Sayantan Chattopadhyay, Aryan Gupta, Satyen Abrol and Archie Gugol	9536
<i>JL-Hate: An Annotated Dataset for Joint Learning of Hate Speech and Target Detection</i> Kaan Büyükdemirci, Izzet Emre Kucukkaya, Eren Ölmez and Cagri Toraman	9543
<i>JMultiWOZ: A Large-Scale Japanese Multi-Domain Task-Oriented Dialogue Dataset</i> Atsumoto Ohashi, Ryu Hirai, Shinya Iizuka and Ryuichiro Higashinaka	9554
<i>Joint Annotation of Morphology and Syntax in Dependency Treebanks</i> Bruno Guillaume, Kim Gerdes, Kirian Guiller, Sylvain Kahane and Yixuan Li	9568
<i>JoTR: A Joint Transformer and Reinforcement Learning Framework for Dialogue Policy Learning</i> Wai-Chung Kwan, Huimin Wang, Hongru Wang, Zezhong Wang, Bin Liang, Xian Wu, Yefeng Zheng and Kam-Fai Wong	9578
<i>JRC-Names-Retrieval: A Standardized Benchmark for Name Search</i> Philip Blair and Kfir Bar	9589
<i>J-SNACS: Adposition and Case Supersenses for Japanese Joshi</i> Tatsuya Aoyama, Chihiro Taguchi and Nathan Schneider	9604
<i>Jump to Conclusions: Short-Cutting Transformers with Linear Transformations</i> Alexander Yom Din, Taelin Karidi, Leshem Choshen and Mor Geva	9615
<i>KazEmoTTS: A Dataset for Kazakh Emotional Text-to-Speech Synthesis</i> Adal Abilbekov, Saida Mussakhojayeva, Rustem Yeshpanov and Huseyin Atakan Varol	9626
<i>KazParC: Kazakh Parallel Corpus for Machine Translation</i> Rustem Yeshpanov, Alina Polonskaya and Huseyin Atakan Varol	9633
<i>KazQAD: Kazakh Open-Domain Question Answering Dataset</i> Rustem Yeshpanov, Pavel Efimov, Leonid Boytsov, Ardak Shalkarbayuli and Pavel Braslavski	9645
<i>KazSAnDRA: Kazakh Sentiment Analysis Dataset of Reviews and Attitudes</i> Rustem Yeshpanov and Huseyin Atakan Varol	9657
<i>KC-GenRe: A Knowledge-constrained Generative Re-ranking Method Based on Large Language Models for Knowledge Graph Completion</i> Yilin Wang, Minghao Hu, Zhen Huang, Dongsheng Li, Dong Yang and Xicheng Lu ..	9668
<i>KCL: Few-shot Named Entity Recognition with Knowledge Graph and Contrastive Learning</i> Shan Zhang, Bin Cao and Jing Fan	9681

<i>KEHRL: Learning Knowledge-Enhanced Language Representations with Hierarchical Reinforcement Learning</i>	
Dongyang Li, Taolin Zhang, Longtao Huang, Chengyu Wang, Xiaofeng He and Hui Xue	9693
<i>KET-QA: A Dataset for Knowledge Enhanced Table Question Answering</i>	
Mengkang Hu, Haoyu Dong, Ping Luo, Shi Han and Dongmei Zhang	9705
<i>Keyphrase Generation: Lessons from a Reproducibility Study</i>	
Edwin Thomas and Sowmya Vajjala	9720
<i>KGConv, a Conversational Corpus Grounded in Wikidata</i>	
Quentin Brabant, Lina M. Rojas Barahona, Gwénolé Lecorvé and Claire Gardent ...	9732
<i>Khan Academy Corpus: A Multilingual Corpus of Khan Academy Lectures</i>	
Dominika Ďurišková, Daniela Jurášová, Matúš Žilinec, Eduard Šubert and Ondřej Bojar	9743
<i>Killkan: The Automatic Speech Recognition Dataset for Kichwa with Morphosyntactic Information</i>	
Chihiro Taguchi, Jefferson Saransig, Dayana Velásquez and David Chiang.....	9753
<i>KIT-19: A Comprehensive Korean Instruction Toolkit on 19 Tasks for Fine-Tuning Korean Large Language Models</i>	
Dongjun Jang, Sungjoo Byun, Hyemi Jo and Hyopil Shin.....	9764
<i>Know-Adapter: Towards Knowledge-Aware Parameter-Efficient Transfer Learning for Few-shot Named Entity Recognition</i>	
Binling Nie, Yiming Shao and Yigang Wang.....	9777
<i>Knowledge-augmented Graph Neural Networks with Concept-aware Attention for Adverse Drug Event Detection</i>	
Ya Gao, Shaoxiong Ji and Pekka Marttinen	9787
<i>Knowledge-aware Attention Network for Medication Effectiveness Prediction</i>	
Yingying Zhang, Xian Wu, Yu Zhang and Yefeng Zheng.....	9799
<i>Knowledge Enhanced Pre-training for Cross-lingual Dense Retrieval</i>	
Hang Zhang, Yeyun Gong, Dayiheng Liu, Shunyu Zhang, Xingwei He, Jiancheng Lv and Jian Guo	9810
<i>Knowledge-enhanced Prompt Tuning for Dialogue-based Relation Extraction with Trigger and Label Semantic</i>	
Hao An, Zhihong Zhu, Xuxin Cheng, Zhiqi Huang and Yuexian Zou	9822
<i>Knowledge GeoGebra: Leveraging Geometry of Relation Embeddings in Knowledge Graph Completion</i>	
Kossi Amouzouvi, Bowen Song, Sahar Vahdati and Jens Lehmann	9832
<i>Knowledge Graphs for Real-World Rumour Verification</i>	
John Dougrez-Lewis, Elena Kochkina, Maria Liakata and Yulan He	9843
<i>Knowledge-Guided Cross-Topic Visual Question Generation</i>	
Hongfei Liu, Guohua Wang, Jiayuan Xie, Jiali Chen, Wenhao Fang and Yi Cai	9854

<i>Knowledge Triplets Derivation from Scientific Publications via Dual-Graph Resonance</i> Kai Zhang, Pengcheng Li, Kaisong Song, Xurui Li, Yangyang Kang, Xuhong Zhang and Xiaozhong Liu.....	9865
<i>KnowVrDU: A Unified Knowledge-aware Prompt-Tuning Framework for Visually-rich Document Understanding</i> Yunqi Zhang, Yubo Chen, Jingzhe Zhu, Jinyu Xu, Shuai Yang, Zhaoliang Wu, Liang Huang, Yongfeng Huang and Shuai Chen	9878
<i>KoCoSa: Korean Context-aware Sarcasm Detection Dataset</i> Yumin Kim, Heejae Suh, Mingi Kim, Dongyeon Won and Hwanhee Lee	9890
<i>KoDialogBench: Evaluating Conversational Understanding of Language Models with Korean Dialogue Benchmark</i> Seongbo Jang, Seonghyeon Lee and Hwanjo Yu	9905
<i>KoFREN: Comprehensive Korean Word Frequency Norms Derived from Large Scale Free Speech Corpora</i> Jin-seo Kim, Anna Seo Gyeong Choi and Sunghye Cho	9926
<i>Konidioms Corpus: A Dataset of Idioms in Konkani Language</i> Naziya Mahamdul Shaikh, Jyoti D. Pawar and Mubarak Banu Sayed	9932
<i>Korean Bio-Medical Corpus (KBMC) for Medical Named Entity Recognition</i> Sungjoo Byun, Jiseung Hong, Sumin Park, Dongjun Jang, Jean Seo, Minseok Kim, Chaeyoung OH and Hyopil Shin	9941
<i>Korean Disaster Safety Information Sign Language Translation Benchmark Dataset</i> Wooyoung Kim, TaeYong Kim, Byeongjin KIM, Myeong Jin MJ Lee, Gitaek Lee, kirok kim, Jisoo Cha and Wooju Kim	9948
<i>Kosmic: Korean Text Similarity Metric Reflecting Honorific Distinctions</i> Yerin Hwang, Yongil Kim, Hyunkyung Bae, Jeesoong Bang, Hwanhee Lee and Kyomin Jung	9954
<i>KPatch: Knowledge Patch to Pre-trained Language Model for Zero-Shot Stance Detection on Social Media</i> Shuohao Lin, Wei Chen, Yunpeng Gao, Zhishu Jiang, Mengqi Liao, Zhiyu Zhang, Shuyuan Zhao and Huaiyu Wan	9961
<i>K-pop Lyric Translation: Dataset, Analysis, and Neural-Modelling</i> Haven Kim, Jongmin Jung, Dasaem Jeong and Juhan Nam.....	9974
<i>L²GC:Lorentzian Linear Graph Convolutional Networks for Node Classification</i> Qiuyu Liang, Weihua Wang, Feilong Bao and Guanglai Gao	9988
<i>Labeling Comic Mischief Content in Online Videos with a Multimodal Hierarchical-Cross-Attention Model</i> Elaheh Baharlouei, Mahsa Shafaei, Yigeng Zhang, Hugo Jair Escalante and Tamar Solorio	9999

<i>Labeling Results of Topic Models: Word Sense Disambiguation as Key Method for Automatic Topic Labeling with GermaNet</i>	
Jennifer Ecker	10014
<i>Landmark-Guided Cross-Speaker Lip Reading with Mutual Information Regularization</i>	
Linzhi Wu, Xingyu Zhang, Yakun Zhang, Changyan Zheng, Tiejun Liu, Liang Xie, Ye Yan and Erwei Yin.....	10023
<i>Language and Speech Technology for Central Kurdish Varieties</i>	
Sina Ahmadi, Daban Jaff, Md Mahfuz Ibn Alam and Antonios Anastasopoulos	10034
<i>Language Models and Semantic Relations: A Dual Relationship</i>	
Olivier Ferret.....	10046
<i>Language Models for Text Classification: Is In-Context Learning Enough?</i>	
Aleksandra Edwards and Jose Camacho-Collados	10058
<i>Language Pivoting from Parallel Corpora for Word Sense Disambiguation of Historical Languages: A Case Study on Latin</i>	
Iacopo Ghinassi, Simone Tedeschi, Paola Marongiu, Roberto Navigli and Barbara McGillivray	10073
<i>Language Technologies as If People Mattered: Centering Communities in Language Technology Development</i>	
Nina Markl, Lauren Hall-Lew and Catherine Lai.....	10085
<i>Language Variety Identification with True Labels</i>	
Marcos Zampieri, Kai North, Tommi Jauhiainen, Mariano Felice, Neha Kumari, Nishant Nair and Yash Mahesh Bangera.....	10100
<i>LANID: LLM-assisted New Intent Discovery</i>	
Lu Fan, Jiashu Pu, Rongsheng Zhang and Xiao-Ming Wu	10110
<i>Large Language Models Are Echo Chambers</i>	
Jan Nehring, Aleksandra Gabryszak, Pascal Jürgens, Aljoscha Burchardt, Stefan Schaffer, Matthias Spielkamp and Birgit Stark.....	10117
<i>Large Language Models as Financial Data Annotators: A Study on Effectiveness and Efficiency</i>	
Toyin D. Aguda, Suchetha Siddagangappa, Elena Kochkina, Simerjot Kaur, Dongsheng Wang and Charese Smiley.....	10124
<i>Large Language Models for Generative Recommendation: A Survey and Visionary Discussions</i>	
Lei Li, Yongfeng Zhang, Dugang Liu and Li Chen.....	10146
<i>Large Language Models Offer an Alternative to the Traditional Approach of Topic Modelling</i>	
Yida Mu, Chun Dong, Kalina Bontcheva and Xingyi Song.....	10160
<i>Latent vs Explicit Knowledge Representation: How ChatGPT Answers Questions about Low-Frequency Entities</i>	
Arianna Graciotti, Valentina Presutti and Rocco Tripodi	10172

<i>LatEval: An Interactive LLMs Evaluation Benchmark with Incomplete Information from Lateral Thinking Puzzles</i>	
Shulin Huang, Shirong Ma, Yinghui Li, Mengzuo Huang, Wuhe Zou, Weidong Zhang and Haitao Zheng	10186
<i>LA-UCL: LLM-Augmented Unsupervised Contrastive Learning Framework for Few-Shot Text Classification</i>	
Jing Zhang, Hui Gao, Peng Zhang, Boda Feng, Wenmin Deng and Yuexian Hou ...	10198
<i>Layer-wise Regularized Dropout for Neural Language Models</i>	
Shiwen Ni, Min Yang, Ruifeng Xu, Chengming Li and Xiping Xiping Hu.....	10208
<i>LayoutLLM: Large Language Model Instruction Tuning for Visually Rich Document Understanding</i>	
Masato Fujitake	10219
<i>LCGbank: A Corpus of Syntactic Analyses Based on Proof Nets</i>	
Aditya Bhargava, Timothy A. D. Fowler and Gerald Penn	10225
<i>LeadEmpathy: An Expert Annotated German Dataset of Empathy in Written Leadership Communication</i>	
Didem Sedefoglu, Allison Claire Lahnala, Jasmin Wagner, Lucie Flek and Sandra Ohly	10237
<i>Learning Bidirectional Morphological Inflection like Humans</i>	
Akiyo Fukatsu, Yuto Harada and Yohei Oseki.....	10249
<i>Learning from Wrong Predictions in Low-Resource Neural Machine Translation</i>	
Jia Cheng Hu, Roberto Cavicchioli, Giulia Berardinelli and Alessandro Capotondi ..	10263
<i>Learning Intrinsic Dimension via Information Bottleneck for Explainable Aspect-based Sentiment Analysis</i>	
Zhenxiao Cheng, Jie Zhou, Wen Wu, Qin Chen and Liang He	10274
<i>Learning Strategies for Robust Argument Mining: An Analysis of Variations in Language and Domain</i>	
Ramon Ruiz-Dolz, Chr-Jr Chiu, Chung-Chi Chen, Noriko Kando and Hsin-Hsi Chen	10286
<i>Lemmatisation of Medieval Greek: Against the Limits of Transformer's Capabilities?</i>	
Colin Swaelens, Pranaydeep Singh, Ilse de Vos and Els Lefever.....	10293
<i>Leros: Learning Explicit Reasoning on Synthesized Data for Commonsense Question Answering</i>	
Chenhao Wang, Pengfei Cao, Jiachun Li, Yubo Chen, Kang Liu, Xiaojian Jiang, Jiexin Xu, Li Qiuxia and Jun Zhao	10303
<i>Lessons from Deploying the First Bilingual Peruvian Sign Language - Spanish Online Dictionary</i>	
Joe Huamani-Malca, Miguel Rodriguez Mondoñedo, Francisco Cerna-Herrera, Gissella Bejarano, Carlos Vásquez Roque, Cesar Augusto Ramos Cantu and Sabina Oporto Pérez	10316

<i>Let's Rectify Step by Step: Improving Aspect-based Sentiment Analysis with Diffusion Models</i> Shunyu Liu, Jie Zhou, Qunxi Zhu, Qin Chen, Qingchun Bai, Jun Xiao and Liang He	10324
<i>Leveraging AMR Graph Structure for Better Sequence-to-Sequence AMR Parsing</i> Linyu Fan, Wu Wu Yiheng, Jun Xie, Junhui Li, Fang Kong and Guodong Zhou	10336
<i>Leveraging Domain Corpora for Enhanced Terminology: The Case of Estonian-English Remote Sensing Termbase</i> Liisi Jakobson, Jelena Kallas and Erko Jakobson	10347
<i>Leveraging Information Redundancy of Real-World Data through Distant Supervision</i> Ariel Cohen, Alexandrine Lanson, Emmanuelle Kempf and Xavier Tannier	10352
<i>Leveraging Linguistically Enhanced Embeddings for Open Information Extraction</i> Fauzan Nayeem Farooqui, Thanmay Jayakumar, Pulkit Mathur and Mansi A. Radke	10365
<i>Leveraging Pre-existing Resources for Data-Efficient Counter-Narrative Generation in Korean</i> Seungyoon Lee, Chanjun Park, DaHyun Jung, Hyeonseok Moon, Jaehyung Seo, Sugyeong Eo and Heuseok Lim	10380
<i>Leveraging Social Context for Humor Recognition and Sense of Humor Evaluation in Social Media with a New Chinese Humor Corpus - HumorWB</i> Zeyuan Zeng, Zefeng Li, Liang Yang and Hongfei Lin	10393
<i>Leveraging Syntactic Dependencies in Disambiguation: The Case of African American English</i> Wilermine Previlon, Alice Rozet, Jotsna Gowda, Bill Dyer, Kevin Tang and Sarah Moeller	10403
<i>Leveraging the Interplay between Syntactic and Acoustic Cues for Optimizing Korean TTS Pause Formation</i> Yejin Jeon, Yunsu Kim and Gary Geunbae Lee	10416
<i>LexAbSumm: Aspect-based Summarization of Legal Decisions</i> Santosh T.Y.S.S., Mahmoud Aly and Matthias Grabmair	10422
<i>LexComSpaL2: A Lexical Complexity Corpus for Spanish as a Foreign Language</i> Jasper Degraeuwe and Patrick Goethals	10432
<i>LexDrafter: Terminology Drafting for Legislative Documents Using Retrieval Augmented Generation</i> Ashish Chouhan and Michael Gertz	10448
<i>LexiVault: A Repository for Psycholinguistic Lexicons of Lesser-studied Languages</i> Hind Saddiki, Samantha Wray and Daisy Li	10459
<i>LFED: A Literary Fiction Evaluation Dataset for Large Language Models</i> Linhao Yu, Qun Liu and Deyi Xiong	10466
<i>LHMKE: A Large-scale Holistic Multi-subject Knowledge Evaluation Benchmark for Chinese Large Language Models</i> Chuang Liu, Renren Jin, Yuqi Ren and Deyi Xiong	10476

<i>LI4: Label-Infused Iterative Information Interacting Based Fact Verification in Question-answering Dialogue</i>	
Xiaocheng Zhang, Chang Wang, Guoping Zhao and Xiaohong Su	10488
<i>LightVLP: A Lightweight Vision-Language Pre-training via Gated Interactive Masked AutoEncoders</i>	
Xingwu Sun, Zhen Yang, Ruobing Xie, Fengzong Lian, Zhanhui Kang and Chengzhong Xu	10499
<i>Limitations of Human Identification of Automatically Generated Text</i>	
Nadège Alavoine, Maximin Coavoux, Emmanuelle Esperança-Rodier, Romane Gallienne, Carlos Gonzalez Gallardo, Jérôme Goulian, Jose G. Moreno, Aurélie Névéol, Didier Schwab, Vincent Segonne and Johanna Simoens	10511
<i>Linear Cross-document Event Coreference Resolution with X-AMR</i>	
Shafiuddin Rehan Ahmed, George Arthur Baker, Evi Judge, Michael Reagan, Kristin Wright-Bettner, Martha Palmer and James H. Martin	10517
<i>LinguaMeta: Unified Metadata for Thousands of Languages</i>	
Sandy Ritchie, Daan van Esch, Uche Okonkwo, Shikhar Vashishth and Emily Drummond	10530
<i>Linguistic Knowledge Can Enhance Encoder-Decoder Models (If You Let It)</i>	
Alessio Miaschi, Felice Dell’Orletta and Giulia Venturi	10539
<i>Linguistic Nudges and Verbal Interaction with Robots, Smart-Speakers, and Humans</i>	
Natalia Kalashnikova, Ioana Vasilescu and Laurence Devillers	10555
<i>Linguistic Rule Induction Improves Adversarial and OOD Robustness in Large Language Models</i>	
Shuoran Jiang, Qingcai Chen, Yang Xiang, Youcheng Pan and Yukang Lin	10565
<i>Linguistic Survey of India and Polyglotta Africana: Two Retrostandardized Digital Editions of Large Historical Collections of Multilingual Wordlists</i>	
Robert Forkel, Johann-Mattis List, Christoph Rzymiski and Guillaume Segerer	10578
<i>Linking Adaptive Structure Induction and Neuron Filtering: A Spectral Perspective for Aspect-based Sentiment Analysis</i>	
Hao Niu, Maoyi Wang, Yun Xiong, Biao Yang, Xing Jia and Zhonglei Guo	10584
<i>Linking Judgement Text to Court Hearing Videos: UK Supreme Court as a Case Study</i>	
Hadeel Saadany, Constantin Orasan, Sophie Walker and Catherine Breslin	10598
<i>Linking Named Entities in Diderot’s Encyclopédie to Wikidata</i>	
Pierre Nugues	10610
<i>Little Red Riding Hood Goes around the Globe: Crosslingual Story Planning and Generation with Large Language Models</i>	
Evgeniia Razumovskaia, Joshua Maynez, Annie Louis, Mirella Lapata and Shashi Narayan	10616
<i>LlamaCare: An Instruction Fine-Tuned Large Language Model for Clinical NLP</i>	
Rumeng Li, Xun Wang and Hong Yu	10632

<i>Llama-VITS: Enhancing TTS Synthesis with Semantic Awareness</i> Xincan Feng and Akifumi Yoshimoto	10642
<i>LLMR: Knowledge Distillation with a Large Language Model-Induced Reward</i> Dongheng Li, Yongchang Hao and Lili Mou	10657
<i>LLMSegm: Surface-level Morphological Segmentation Using Large Language Model</i> Marko Pranjic, Marko Robnik-Sikonja and Senja Pollak	10665
<i>LM-Combiner: A Contextual Rewriting Model for Chinese Grammatical Error Correction</i> Yixuan Wang, Baoxin Wang, Yijun Liu, Dayong Wu and Wanxiang Che	10675
<i>Locally Differentially Private In-Context Learning</i> Chunyan Zheng, Keke Sun, Wenhao Zhao, Haibo Zhou, Lixing Jiang, Shaoyang Song and Chunlai Zhou	10686
<i>LocalTweets to LocalHealth: A Mental Health Surveillance Framework Based on Twitter Data</i> Vijeta Deshpande, Minhwa Lee, Zonghai Yao, Zihao Zhang, Jason Brian Gibbons and Hong Yu	10698
<i>Loflòc: A Morphological Lexicon for Occitan using Universal Dependencies</i> Marianne Vergez-Couret, Myriam Bras, Aleksandra Miletic and Clamença Poujade	10716
<i>Logging Keystrokes in Writing by English Learners</i> Georgios Velentzas, Andrew Caines, Rita Borgo, Erin Pacquetet, Clive Hamilton, Taylor Arnold, Diane Nicholls, Paula Buttery, Thomas Gaillat, Nicolas Ballier and Helen Yannakoudakis	10725
<i>Logic Rules as Explanations for Legal Case Retrieval</i> ZhongXiang Sun, Kepu Zhang, Weijie Yu, Haoyu Wang and Jun Xu	10747
<i>LoNAS: Elastic Low-Rank Adapters for Efficient Large Language Models</i> Juan Pablo Munoz, Jinjie Yuan, Yi Zheng and Nilesh Jain	10760
<i>LongDocFACTScore: Evaluating the Factuality of Long Document Abstractive Summarisation</i> Jennifer A. Bishop, Sophia Ananiadou and Qianqian Xie	10777
<i>Longform Multimodal Lay Summarization of Scientific Papers: Towards Automatically Generating Science Blogs from Research Articles</i> Sandeep Kumar, Guneet Singh Kohli, Tirthankar Ghosal and Asif Ekbal	10790
<i>Look before You Leap: Dual Logical Verification for Knowledge-based Visual Question Generation</i> Xumeng Liu, Wenya Guo, Ying Zhang, Xubo Liu, Yu Zhao, Shenglong Yu and Xiaojie Yuan	10802
<i>LoSST-AD: A Longitudinal Corpus for Tracking Alzheimer's Disease Related Changes in Spontaneous Speech</i> Ulla Petti and Anna Korhonen	10813
<i>Low-Rank Prune-And-Factorize for Language Model Compression</i> Siyu Ren and Kenny Q. Zhu	10822

<i>M2SA: Multimodal and Multilingual Model for Sentiment Analysis of Tweets</i> Gaurish Thakkar, Sherzod Hakimov and Marko Tadić	10833
<i>M3: A Multi-Task Mixed-Objective Learning Framework for Open-Domain Multi-Hop Dense Sentence Retrieval</i> Yang Bai, Anthony Colas, Christan Grant and Zhe Wang	10846
<i>m3P: Towards Multimodal Multilingual Translation with Multimodal Prompt</i> Jian Yang, Hongcheng Guo, Yuwei Yin, Jiaqi Bai, Bing Wang, Jiaheng Liu, Xinnian Liang, LinZheng Chai, Liqun Yang and Zhoujun Li	10858
<i>M3TCM: Multi-modal Multi-task Context Model for Utterance Classification in Motivational Interviews</i> Sayed Muddashir Hossain, Jan Alexandersson and Philipp Müller	10872
<i>MaCmS: Magahi Code-mixed Dataset for Sentiment Analysis</i> Priya Rani, Theodorus Fransen, John P. McCrae and Gaurav Negi	10880
<i>MAGIC: Multi-Argument Generation with Self-Refinement for Domain Generalization in Automatic Fact-Checking</i> Wei-Yu Kao and An-Zi Yen	10891
<i>MAGPIE: Multi-Task Analysis of Media-Bias Generalization with Pre-Trained Identification of Expressions</i> Tomáš Horych, Martin Paul Wessel, Jan Philip Wahle, Terry Ruas, Jerome Waßmuth, André Greiner-Petter, Akiko Aizawa, Bela Gipp and Timo Spinde	10903
<i>MaiBaam: A Multi-Dialectal Bavarian Universal Dependency Treebank</i> Verena Blaschke, Barbara Kovačić, Siyao Peng, Hinrich Schütze and Barbara Plank	10921
<i>MaintIE: A Fine-Grained Annotation Schema and Benchmark for Information Extraction from Maintenance Short Texts</i> Tyler K. Bikaun, Tim French, Michael Stewart, Wei Liu and Melinda Hodkiewicz	10939
<i>Majority Rules Guided Aspect-Category Based Sentiment Analysis via Label Prior Knowledge</i> Lin Li, Shaopeng Tang and Renwei Wu	10952
<i>Make Prompt-based Black-Box Tuning Colorful: Boosting Model Generalization from Three Orthogonal Perspectives</i> Qiushi Sun, Chengcheng Han, Nuo Chen, Renyu Zhu, Jingyang Gong, Xiang Li and Ming Gao	10958
<i>Making Pre-trained Language Models Better Continual Few-Shot Relation Extractors</i> Shengkun Ma, Jiale Han, Yi Liang and Bo Cheng	10970
<i>Making Sentence Embeddings Robust to User-Generated Content</i> Lydia Nishimwe, Benoît Sagot and Rachel Bawden	10984
<i>Malaysian English News Decoded: A Linguistic Resource for Named Entity and Relation Extraction</i> MohanRaj Chanthran, Lay-Ki Soon, Huey Fang Ong and Bhawani Selvaretnam	10999
<i>mALBERT: Is a Compact Multilingual BERT Model Still Worth It?</i> Christophe Servan, Sahar Ghannay and Sophie Rosset	11023

<i>ManNER & ManPOS: Pioneering NLP for Endangered Manchu Language</i> Sangah Lee, Sungjoo Byun, Jean Seo and Minha Kang	11030
<i>Mapping the Past: Geographically Linking an Early 20th Century Swedish Encyclopedia with Wikidata</i> Axel Ahlin, Alfred Myrne Blåder and Pierre Nugues	11040
<i>Mapping Work Task Descriptions from German Job Ads on the O*NET Work Activities Ontology</i> Ann-Sophie Gnehm and Simon Clematide.....	11049
<i>MARASTA: A Multi-dialectal Arabic Cross-domain Stance Corpus</i> Anis Charfi, Mabrouka Ben-Sghaier, Andria Samy Raouf Atalla, Raghda Akasheh, Sara Al-Emadi and Wajdi Zaghouni	11060
<i>Massively Multilingual Token-Based Typology Using the Parallel Bible Corpus</i> Amanda Kann.....	11070
<i>Mathematical Entities: Corpora and Benchmarks</i> Jacob Collard, Valeria de Paiva and Eswaran Subrahmanian	11080
<i>MccSTN: Multi-Scale Contrast and Fine-Grained Feature Fusion Networks for Subject-driven Style Transfer</i> Honggang Zhao, Chunling Xiao, Jiayi Yang, Guozhu Jin and Mingyong Li.....	11090
<i>MCIL: Multimodal Counterfactual Instance Learning for Low-resource Entity-based Multimodal Information Extraction</i> Baohang Zhou, Ying Zhang, Kehui Song, Hongru Wang, Yu Zhao, Xuhui Sui and Xiaojie Yuan.....	11101
<i>MCTS: A Multi-Reference Chinese Text Simplification Dataset</i> Ruining Chong, Luming Lu, Liner Yang, Jinran Nie, Zhenghao Liu, Shuo Wang, Shuhan Zhou, Yaoxin Li and Erhong Yang	11111
<i>MDS: A Fine-Grained Dataset for Multi-Modal Dialogue Summarization</i> Zhipeng Liu, Xiaoming Zhang, Litian Zhang and Zelong Yu	11123
<i>Measuring Cross-Text Cohesion for Segmentation Similarity Scoring</i> Gerardo Ocampo Diaz and Jessica Ouyang	11138
<i>Medical Entity Disambiguation with Medical Mention Relation and Fine-grained Entity Knowledge</i> Wenpeng Lu, Guobiao Zhang, Xueping Peng, Hongjiao Guan and Shoujin Wang ..	11148
<i>Medical Vision-Language Pre-Training for Brain Abnormalities</i> Masoud Monajatipoor, Zi-Yi Dou, Aichi Chien, Nanyun Peng and Kai-Wei Chang ...	11159
<i>MedMT5: An Open-Source Multilingual Text-to-Text LLM for the Medical Domain</i> Iker García-Ferrero, Rodrigo Agerri, Aitziber Atutxa Salazar, Elena Cabrio, Iker de la Iglesia, Alberto Lavelli, Bernardo Magnini, Benjamin Molinet, Johana Ramirez-Romero, German Rigau, Jose Maria Villa-Gonzalez, Serena Villata and Andrea Zaninello.....	11165
<i>MedQA-SWE - a Clinical Question & Answer Dataset for Swedish</i> Niclas Hertzberg and Anna Lokrantz	11178

<i>MemoryPrompt: A Light Wrapper to Improve Context Tracking in Pre-trained Language Models</i> Nathanael Carraz Rakotonirina and Marco Baroni	11187
<i>MentalHelp: A Multi-Task Dataset for Mental Health in Social Media</i> Nishat Raihan, Sadiya Sayara Chowdhury Puspo, Shafkat Farabi, Ana-Maria Bucur, Tharindu Ranasinghe and Marcos Zampieri.....	11196
<i>MentalRiskES: A New Corpus for Early Detection of Mental Disorders in Spanish</i> Alba M. Mármol Romero, Adrián Moreno Muñoz, Flor Miriam Plaza-del-Arco, M. Dolores Molina González, María Teresa Martín Valdivia, L. Alfonso Ureña-López and Arturo Montejó Ráez	11204
<i>Meta-Adapter for Self-Supervised Speech Models: A Solution to Low-Resource Speech Recognition Challenges</i> Yaqi Chen, Hao Zhang, Xukui Yang, Wenlin Zhang and Dan Qu	11215
<i>Meta-Cognitive Analysis: Evaluating Declarative and Procedural Knowledge in Datasets and Large Language Models</i> Zhuoqun Li, Hongyu Lin, Yaojie Lu, Hao Xiang, Xianpei Han and Le Sun	11222
<i>Meta-Evaluation of Sentence Simplification Metrics</i> Noof Abdullah Alfeear, Dimitar Kazakov and Hend Al-Khalifa.....	11229
<i>Metaphors in Online Religious Communication: A Detailed Dataset and Cross-Genre Metaphor Detection</i> Sebastian Reimann and Tatjana Scheffler	11236
<i>MEVTR: A Multilingual Model Enhanced with Visual Text Representations</i> Xiaohua Wang, Wenlong Fei, Min Hu, Qingyu Zhang and Aoqiang Zhu	11247
<i>mForms : Multimodal Form Filling with Question Answering</i> Larry Heck, Simon Heck and Anirudh S. Sundar.....	11262
<i>MHGRL: An Effective Representation Learning Model for Electronic Health Records</i> Feiyan Liu, Liangzhi Li, Xiaoli Wang, Feng Luo, Chang Liu, Jinsong Su and Yiming Qian 11272	
<i>MiDe22: An Annotated Multi-Event Tweet Dataset for Misinformation Detection</i> Cagri Toraman, Oguzhan Ozcelik, Furkan Sahinuc and Fazli Can.....	11283
<i>Mind Your Neighbours: Leveraging Analogous Instances for Rhetorical Role Labeling for Legal Documents</i> Santosh T.Y.S.S., Hassan Sarwat, Ahmed Mohamed Abdelaal Abdou and Matthias Grab- mair	11296
<i>MinT: Boosting Generalization in Mathematical Reasoning via Multi-view Fine-tuning</i> Zhenwen Liang, Dian Yu, Xiaoman Pan, Wenlin Yao, Qingkai Zeng, Xiangliang Zhang and Dong Yu	11307
<i>Mitigating Linguistic Artifacts in Emotion Recognition for Conversations from TV Scripts to Daily Conversations</i> Donovan Ong, Shuo Sun, Jian Su and Bin Chen	11319

<i>Mitigating Misleading Chain-of-Thought Reasoning with Selective Filtering</i> Yexin Wu, Zhuosheng Zhang and Hai Zhao	11325
<i>Mitigating Shortcuts in Language Models with Soft Label Encoding</i> Zirui He, Huiqi Deng, Haiyan Zhao, Ninghao Liu and Mengnan Du	11341
<i>Mitigating Translationese in Low-resource Languages: The Storyboard Approach</i> Garry Kuwanto, Eno-Abasi E. Urua, Priscilla Amondi Amuok, Shamsuddeen Hassan Muham- mad, Anuoluwapo Aremu, Verrah Otiende, Loice Emma Nanyanga, Teresiah W. Nyoike, Aniefon D. Akpan, Nsima Ab Udouboh, Idongesit Udeme Archibong, Idara Effiong Moses, Ifeoluwatayo A. Ige, Benjamin Ajibade, Olumide Benjamin Awokoya, Idris Abdulmumin, Saminu Mohammad Aliyu, Ruqayya Nasir Iro, Ibrahim Said Ahmad, Deontae Smith, Praise-EL Michaels, David Ife- oluwa Adelani, Derry Tanti Wijaya and Anietie Andy	11349
<i>MixRED: A Mix-lingual Relation Extraction Dataset</i> Lingxing Kong, Yougang Chu, Zheng Ma, Jianbing Zhang, Liang He and Jiajun Chen	11361
<i>Mixture-of-LoRAs: An Efficient Multitask Tuning Method for Large Language Models</i> Wenfeng Feng, Chuzhan Hao, Yuewei Zhang, Yu Han and Hao Wang	11371
<i>Mixture-of-Prompt-Experts for Multi-modal Semantic Understanding</i> Zichen Wu, Hsiu-Yuan Huang, Fanyi Qu and Yunfang Wu	11381
<i>MKeCL: Medical Knowledge-Enhanced Contrastive Learning for Few-shot Disease Diagnosis</i> Yutian Zhao, Huimin Wang, Xian Wu and Yefeng Zheng	11394
<i>MLDSP-MA: Multidimensional Attention for Multi-Round Long Dialogue Sentiment Prediction</i> Yunfei Yin, Congrui Zou, Zheng Yuan and Xianjian Bao	11405
<i>MMAD: Multi-modal Movie Audio Description</i> Xiaojun Ye, Junhao Chen, Xiang Li, Haidong Xin, Chao Li, Sheng Zhou and Jiajun Bu	11415
<i>MMAPS: End-to-End Multi-Grained Multi-Modal Attribute-Aware Product Summarization</i> Tao Chen, Ze Lin, Hui Li, Jiayi Ji, Yiyi Zhou, Guanbin Li and Rongrong Ji	11429
<i>MM-IGLU: Multi-Modal Interactive Grounded Language Understanding</i> Claudiu Daniel Hromei, Daniele Margiotta, Danilo Croce and Roberto Basili	11440
<i>MNER-MI: A Multi-image Dataset for Multimodal Named Entity Recognition in Social Media</i> Shizhou Huang, Bo Xu, Changqun Li, Jiabo Ye and xin Lin	11452
<i>Modalities Should Be Appropriately Leveraged: Uncertainty Guidance for Multimodal Chinese Spelling Correction</i> Yongliang Lin, Zhen Zhang, Mengting Hu, Yufei Sun and Yuzhi Zhang	11463
<i>MoDE-CoTD: Chain-of-Thought Distillation for Complex Reasoning Tasks with Mixture of De- coupled LoRA-Experts</i> Xiang Li, Shizhu He, Jiayu Wu, Zhao Yang, Yao Xu, Yang jun Jun, Haifeng Liu, Kang Liu and Jun Zhao	11475
<i>Model-Agnostic Cross-Lingual Training for Discourse Representation Structure Parsing</i> Jiangming Liu	11486

<i>Modeling Low-Resource Health Coaching Dialogues via Neuro-Symbolic Goal Summarization and Text-Units-Text Generation</i>	
Yue Zhou, Barbara Di Eugenio, Brian Ziebart, Lisa Sharp, Bing Liu and Nikolaos Agadakos	11498
<i>Modeling Orthographic Variation Improves NLP Performance for Nigerian Pidgin</i>	
Pin-Jie Lin, Merel Scholman, Muhammed Saeed and Vera Demberg	11510
<i>Modeling the Quality of Dialogical Explanations</i>	
Milad Alshomary, Felix Lange, Meisam Booshehri, Meghdut Sengupta, Philipp Cimiano and Henning Wachsmuth	11523
<i>Modelling and Linking an Old Latin-Portuguese Dictionary to the LiLa Knowledge Base</i>	
Lucas Consolin Dezotti, Marco Passarotti and Francesco Mambri	11537
<i>Modelling Argumentation for an User Opinion Aggregation Tool</i>	
Pablo Weingart, Thiemo Wambsganss and Matthias Soellner	11548
<i>MoNMT: Modularly Leveraging Monolingual and Bilingual Knowledge for Neural Machine Translation</i>	
Jianhui Pang, Baosong Yang, Derek F. Wong, Dayiheng Liu, Xiangpeng Wei, Jun Xie and Lidia S. Chao	11560
<i>Monolingual Paraphrase Detection Corpus for Low Resource Pashto Language at Sentence Level</i>	
Iqra Ali, Hidetaka Kamigaito and Taro Watanabe	11574
<i>MoPE: Mixture of Prefix Experts for Zero-Shot Dialogue State Tracking</i>	
Tianwen Tang, Tong Zhu, Haodong Liu, Yin Bai, Jia Cheng and Wenliang Chen	11582
<i>MORE-3S: Multimodal-based Offline Reinforcement Learning with Shared Semantic Spaces</i>	
Tianyu Zheng, Ge Zhang, Xingwei Qu, Ming Kuang, Wenhao Huang and Zhaofeng He	11593
<i>Morpheme Sense Disambiguation: A New Task Aiming for Understanding the Language at Character Level</i>	
Yue Wang, Hua Zheng, Yaqi Yin, Hansi Wang, Qiliang Liang and Yang Liu	11605
<i>Motion Capture Analysis of Verb and Adjective Types in Austrian Sign Language (ÖGS)</i>	
Julia Krebs, Evguenia A. Malaia, Isabella Fessl, Hans-Peter Wiesinger, Dietmar Roehm, Ronnie Wilbur and Hermann Schwameder	11619
<i>Motion Generation from Fine-grained Textual Descriptions</i>	
Kunhang Li and Yansong Feng	11625
<i>Motivational Interviewing Transcripts Annotated with Global Scores</i>	
Ben Cohen, Moreah Zisquit, Stav Yosef, Doron Friedman and Kfir Bar	11642
<i>MoZIP: A Multilingual Benchmark to Evaluate Large Language Models in Intellectual Property</i>	
Shiwen Ni, Minghuan Tan, Yuelin Bai, Fuqiang Niu, Min Yang, Bowen Zhang, Ruifeng Xu, Xiaojun Chen, Chengming Li and Xiping Hu	11658
<i>MRC-based Nested Medical NER with Co-prediction and Adaptive Pre-training</i>	
Xiaojing Du, hanjie Zhao, Danyan Xing, Yuxiang Jia and Hongying Zan	11669

<i>MRT: Multi-modal Short- and Long-range Temporal Convolutional Network for Time-sync Comment Video Behavior Prediction</i>	
Weihao Zhao, Weidong He, Hao Wang, Haoyang Bi, Han Wu, Chen Zhu, Tong Xu and Enhong Chen.....	11680
<i>MUCH: A Multimodal Corpus Construction for Conversational Humor Recognition Based on Chinese Sitcom</i>	
Hongyu Guo, Wenbo Shang, Xueyao Zhang and Binyang Li.....	11692
<i>Multi-Channel Spatio-Temporal Transformer for Sign Language Production</i>	
Xiaohan Ma, Rize Jin and Tae-Sun Chung.....	11699
<i>MULTICOLLAB: A Multimodal Corpus of Dialogues for Analyzing Collaboration and Frustration in Language</i>	
Michael Peechatt, Cecilia Ovesdotter Alm and Reynold Bailey.....	11713
<i>Multi-Dimensional Machine Translation Evaluation: Model Evaluation and Resource for Korean</i>	
Dojun Park and Sebastian Padó.....	11723
<i>Multi-domain Hate Speech Detection Using Dual Contrastive Learning and Paralinguistic Features</i>	
Somaiyeh Dehghan and Berrin Yanıkoğlu.....	11745
<i>Multi-Grained Conversational Graph Network for Retrieval-based Dialogue Systems</i>	
Quan Tu, Chongyang Tao and Rui Yan.....	11756
<i>Multi-Granularity Fusion Text Semantic Matching Based on WoBERT</i>	
Hongchun Yu, Wei Pan, Xing Fan and Hanqi Li.....	11766
<i>MultiLeg: Dataset for Text Sanitisation in Less-resourced Languages</i>	
Rinalds Viksna and Inguna Skadiņa.....	11776
<i>MultiLexBATS: Multilingual Dataset of Lexical Semantic Relations</i>	
Dagmar Gromann, Hugo Goncalo Oliveira, Lucia Pitarch, Elena-Simona Apostol, Jordi Bernad, Eliot Bytyçi, Chiara Cantone, Sara Carvalho, Francesca Frontini, Radovan Garabik, Jorge Gracia, Letizia Granata, Fahad Khan, Timotej Knez, Penny Labropoulou, Chaya Liebeskind, Maria Pia Di Buono, Ana Ostroški Anić, Sigita Rackevičienė, Ricardo Rodrigues, Gilles Sérasset, Linas Selmistraitis, Mahammadou Sidibé, Purificação Silvano, Blerina Spahiu, Enriketa Sogutlu, Ranka Stanković, Ciprian-Octavian Truică, Giedre Valunaite Oleskeviciene, Slavko Zitnik and Katerina Zdravkova.....	11783
<i>Multilingual Brain Surgeon: Large Language Models Can Be Compressed Leaving No Language behind</i>	
Hongchuan Zeng, Hongshen Xu, Lu Chen and Kai Yu.....	11794
<i>Multilingual Coreference Resolution in Low-resource South Asian Languages</i>	
Ritwik Mishra, Pooja Desur, Rajiv Ratn Shah and Ponnurangam Kumaraguru.....	11813
<i>Multilingual Generation in Abstractive Summarization: A Comparative Study</i>	
Jinpeng Li, Jiase Chen, Huadong Chen, Dongyan Zhao and Rui Yan.....	11827
<i>Multilinguality or Back-translation? A Case Study with Estonian</i>	
Elizaveta Korotkova, Taido Purason, Agnes Luhtaru and Mark Fishel.....	11838

<i>Multilingual Sentence-T5: Scalable Sentence Encoders for Multilingual Applications</i> Chihiro Yano, Akihiko Fukuchi, Shoko Fukasawa, Hideyuki Tachibana and Yotaro Watanabe	11849
<i>Multilingual Substitution-based Word Sense Induction</i> Denis Kokosinskii and Nikolay Arefyev	11859
<i>Multilingual Turn-taking Prediction Using Voice Activity Projection</i> Koji Inoue, Bing'er Jiang, Erik Ekstedt, Tatsuya Kawahara and Gabriel Skantze	11873
<i>Multimodal and Multilingual Laughter Detection in Stand-Up Comedy Videos</i> Anna Kuznetsova and Carlo Strapparava	11884
<i>Multimodal Behaviour in an Online Environment: The GEHM Zoom Corpus Collection</i> Patrizia Paggio, Manex Agirrezabal, Costanza Navarretta and Leo Vitasovic	11890
<i>Multimodal Cross-Document Event Coreference Resolution Using Linear Semantic Transfer and Mixed-Modality Ensembles</i> Abhijnan Nath, Huma Jamil, Shafiuddin Rehan Ahmed, George Arthur Baker, Rahul Ghosh, James H. Martin, Nathaniel Blanchard and Nikhil Krishnaswamy	11901
<i>Multimodal Cross-lingual Phrase Retrieval</i> Chuanqi Dong, Wenjie Zhou, Xiangyu Duan, Yuqi Zhang and Min Zhang	11917
<i>Multimodal Language Models Show Evidence of Embodied Simulation</i> Cameron R. Jones and Sean Trott	11928
<i>Multi-modal Semantic Understanding with Contrastive Cross-modal Feature Alignment</i> Ming Zhang, Ke Chang and Yunfang Wu	11934
<i>Multi-Objective Forward Reasoning and Multi-Reward Backward Refinement for Product Review Summarization</i> Libo Sun, Siyuan Wang, Meng Han, Ruofei Lai, Xinyu Zhang, Xuanjing Huang and Zhongyu Wei	11944
<i>Multi-perspective Improvement of Knowledge Graph Completion with Large Language Models</i> Derong Xu, Ziheng Zhang, Zhenxi Lin, Xian Wu, Zhihong Zhu, Tong Xu, Xiangyu Zhao, Yefeng Zheng and Enhong Chen	11956
<i>Multi-Stage Multi-Modal Pre-Training for Automatic Speech Recognition</i> Yash Jain, David M. Chan, Pranav Dheram, Aparna Khare, Olabanji Shonibare, Venkatesh Ravichandran and Shalini Ghosh	11969
<i>Multi-stream Information Fusion Framework for Emotional Support Conversation</i> Yinan Bao, Dou Hu, Lingwei Wei, Shuchong Wei, Wei Zhou and Songlin Hu	11981
<i>Multi-Tiered Cantonese Word Segmentation</i> Charles Lam, Chaak-ming Lau and Jackson L. Lee	11993
<i>Murre24: Dialect Identification of Finnish Internet Forum Messages</i> Olli Kuparinen	12003
<i>MVP: Minimal Viable Phrase for Long Text Understanding</i> Louis Clouatre, Amal Zouaq and Sarath Chandar	12016

<i>MWE-Finder: A Demonstration</i>	
Jan Odijk, Martin Kroon, Tijmen Baarda, Ben Bonfil and Sheean Spoel	12027
<i>myMediCon: End-to-End Burmese Automatic Speech Recognition for Medical Conversations</i>	
Hay Man Htun, Ye Kyaw Thu, Hutchatai Chanlekha, Kotaro Funakoshi and Thepchai Supnithi	12032
<i>My Science Tutor (MyST)—a Large Corpus of Children’s Conversational Speech</i>	
Sameer Pradhan, Ronald A. Cole and Wayne H. Ward	12040
<i>NAIST-SIC-Aligned: An Aligned English-Japanese Simultaneous Interpretation Corpus</i>	
Jinming Zhao, Katsuhito Sudoh, Satoshi Nakamura, Yuka Ko, Kosuke Doi and Ryo Fukuda	12046
<i>NarrativeTime: Dense Temporal Annotation on a Timeline</i>	
Anna Rogers, Marzena Karpinska, Ankita Gupta, Vladislav Lialin, Gregory Smelkov and Anna Rumshisky	12053
<i>Navigating Prompt Complexity for Zero-Shot Classification: A Study of Large Language Models in Computational Social Science</i>	
Yida Mu, Ben P. Wu, William Thorne, Ambrose Robinson, Nikolaos Aletras, Carolina Scarton, Kalina Bontcheva and Xingyi Song	12074
<i>NB Uttale: A Norwegian Pronunciation Lexicon with Dialect Variation</i>	
Marie Iversdatter Røsok and Ingerid Løyning Dale	12087
<i>Negation Scope Conversion: Towards a Unified Negation-Annotated Dataset</i>	
Asahi Yoshida, Yoshihide Kato and Shigeki Matsubara	12093
<i>Negation Triplet Extraction with Syntactic Dependency and Semantic Consistency</i>	
Yuchen Shi, Deqing Yang, Jingping Liu, Yanghua Xiao, Zongyu Wang and Huimin Xu	12100
<i>nEMO: Dataset of Emotional Speech in Polish</i>	
Iwona Christop	12111
<i>NER-guided Comprehensive Hierarchy-aware Prompt Tuning for Hierarchical Text Classification</i>	
Fuhan Cai, Duo Liu, Zhongqiang Zhang, Ge Liu, Xiaozhe Yang and Xiangzhong Fang	12117
<i>Nested Event Extraction upon Pivot Element Recognition</i>	
Weicheng Ren, Zixuan Li, Xiaolong Jin, Long Bai, Miao Su, Yantao Liu, Saiping Guan, Jiafeng Guo and Xueqi Cheng	12127
<i>Nested Noun Phrase Identification Using BERT</i>	
Shweta Misra and Johan Boye	12138
<i>Neural Machine Translation between Low-Resource Languages with Synthetic Pivoting</i>	
Khalid Ahmed and Jan Buys	12144
<i>Neural Multimodal Topic Modeling: A Comprehensive Evaluation</i>	
Felipe Gonzalez-Pizarro and Giuseppe Carenini	12159
<i>New Datasets for Automatic Detection of Textual Entailment and of Contradictions between Sentences in French</i>	
Maximos Skandalis, Richard Moot, Christian Retoré and Simon Robillard	12173

<i>New Evaluation Methodology for Qualitatively Comparing Classification Models</i> Ahmad Aljanaideh	12187
<i>New Intent Discovery with Attracting and Dispersing Prototype</i> Shun Zhang, Jian Yang, Jiaqi Bai, Chaoran Yan, Tongliang Li, Zhao Yan and Zhoujun Li	12193
<i>New Methods for Exploring Intonosyntax: Introducing an Intonosyntactic Treebank for Nigerian Pidgin</i> Emmett Strickland, Anne Lacheret-Dujour, Sylvain Kahane, Marc Evrard, Perrine Quennehen, Bernard Caron, Francis Egbokhare and Bruno Guillaume	12207
<i>New Proposal of Greenberg's Universal 14 from Typometrics</i> Antoni Brosa-Rodríguez and Sylvain Kahane	12217
<i>New Semantic Task for the French Spoken Language Understanding MEDIA Benchmark</i> Nadège Alavoine, Gaëlle Laperrière, Christophe Servan, Sahar Ghannay and Sophie Rosset	12227
<i>NGLUEni: Benchmarking and Adapting Pretrained Language Models for Nguni Languages</i> Francois Meyer, Haiyue Song, Abhisek Chakrabarty, Jan Buys, Raj Dabre and Hideki Tanaka	12247
<i>NLoPT: N-gram Enhanced Low-Rank Task Adaptive Pre-training for Efficient Language Model Adaption</i> Hao Gu, Jiangyan Yi, Zheng Lian, Jianhua Tao and Xinrui Yan	12259
<i>NLPre: A Revised Approach towards Language-centric Benchmarking of Natural Language Preprocessing Systems</i> Martyna Wiącek, Piotr Rybak, Łukasz Pszenny and Alina Wróblewska	12271
<i>No Need for Large-Scale Search: Exploring Large Language Models in Complex Knowledge Base Question Answering</i> Shouhui Wang and Biao Qin	12288
<i>Non-Essential Is NEcessary: Order-agnostic Multi-hop Question Generation</i> Kyungho Kim, Seongmin Park, Junseo Lee and Jihwa Lee	12300
<i>NSina: A News Corpus for Sinhala</i> Hansi Hettiarachchi, Damith Premasiri, Lasitha Randunu Chandrakantha Uyangodage and Tharindu Ranasinghe	12307
<i>Null Subjects in Spanish as a Machine Translation Problem</i> Jose Diego Suarez and Luis Chiruzzo	12313
<i>NumHG: A Dataset for Number-Focused Headline Generation</i> Jian-Tao Huang, Chung-Chi Chen, Hen-Hsen Huang and Hsin-Hsi Chen	12323
<i>NutFrame: Frame-based Conceptual Structure Induction with LLMs</i> Shaoru Guo, Yubo Chen, Kang Liu, Ru Li and Jun Zhao	12330

<i>OATS: A Challenge Dataset for Opinion Aspect Target Sentiment Joint Detection for Aspect-Based Sentiment Analysis</i>	
Siva Uday Sampreeth Chebolu, Franck Deroncourt, Nedim Lipka and Tamar Solorio	12336
<i>OLViT: Multi-Modal State Tracking via Attention-Based Embeddings for Video-Grounded Dialog</i>	
Adnen Abdessaied, Manuel Hochmeister and Andreas Bulling	12348
<i>On an Intermediate Task for Classifying URL Citations on Scholarly Papers</i>	
Kazuhiro Wada, Masaya Tsunokake and Shigeki Matsubara	12359
<i>On Leveraging Encoder-only Pre-trained Language Models for Effective Keyphrase Generation</i>	
Di Wu, Wasi U. Ahmad and Kai-Wei Chang	12370
<i>On Modelling Corpus Citations in Computational Lexical Resources</i>	
Fahad Khan, Maxim Ionov, Christian Chiarcos, Laurent Romary, Gilles Sérasset and Besim Kabashi	12385
<i>On the Adaptation of Unlimiformer for Decoder-Only Transformers</i>	
Kian Ahrabian, Alon Benhaim, Barun Patra, Jay Pujara, Saksham Singhal and Xia Song	12395
<i>On the Relationship between Skill Neurons and Robustness in Prompt Tuning</i>	
Leon Ackermann and Xenia Isabel Ohmer	12403
<i>On the Scaling Laws of Geographical Representation in Language Models</i>	
Nathan Godey, Éric de la Clergerie and Benoît Sagot	12416
<i>On the Use of Silver Standard Data for Zero-shot Classification Tasks in Information Extraction</i>	
Jianwei Wang, Tianyin Wang and Ziqian Zeng	12423
<i>On the Way to Lossless Compression of Language Transformers: Exploring Cross-Domain Properties of Quantization</i>	
Nikita Martynov, Aleksei Goncharov, Gleb Kumichev, Evgeniy Egorov, Stanislav Vladimirovich Pavlov, Mikhail Sergeevich Durinov, Aleksandr Sergeevich Zuev and Egor Anatolievich Filimonov	12435
<i>On Zero-Shot Counterspeech Generation by LLMs</i>	
Punyajoy Saha, Aalok Agrawal, Abhik Jana, Chris Biemann and Animesh Mukherjee	12443
<i>OOVs in the Spotlight: How to Inflect Them?</i>	
Tomáš Sourada, Jana Straková and Rudolf Rosa	12455
<i>OpenMSD: Towards Multilingual Scientific Documents Similarity Measurement</i>	
Yang Gao, Ji Ma, Ivan Korotkov, Keith Hall, Dana Alon and Donald Metzler	12467
<i>Opinion Mining Using Pre-Trained Large Language Models: Identifying the Type, Polarity, Intensity, Expression, and Source of Private States</i>	
Saeed Ahmadnia, Arash Yousefi Jordehi, Mahsa Hosseini Khasheh Heyran, SeyedAbolghasem Mirroshandel and Owen Rambow	12481
<i>Opinions Are Not Always Positive: Debiasing Opinion Summarization with Model-Specific and Model-Agnostic Methods</i>	
Yanyue Zhang, Yilong Lai, Zhenglin Wang, Pengfei Li, Deyu Zhou and Yulan He	12496

<i>Optimizing Language Augmentation for Multilingual Large Language Models: A Case Study on Korean</i>	
ChangSu Choi, Yongbin Jeong, Seoyoon Park, Inho Won, HyeonSeok Lim, SangMin Kim, Yejee Kang, Chanhyuk Yoon, Jaewan Park, Yiseul Lee, HyeJin Lee, Younggyun Hahm, Hansaem Kim and KyungTae Lim	12514
<i>ORTicket: Let One Robust BERT Ticket Transfer across Different Tasks</i>	
Yuhao Zhou, Wenxiang Chen, Rui Zheng, Zhiheng Xi, Tao Gui, Qi Zhang and Xuanjing Huang	12527
<i>Out-of-Domain Intent Detection Considering Multi-Turn Dialogue Contexts</i>	
Hao Lang, Yinhe Zheng, Binyuan Hui, Fei Huang and Yongbin Li	12539
<i>Out of the Mouths of MPs: Speaker Attribution in Parliamentary Debates</i>	
Ines Rehbein, Josef Ruppenhofer, Annelen Brunner and Simone Paolo Ponzetto ..	12553
<i>PACAR: Automated Fact-Checking with Planning and Customized Action Reasoning Using Large Language Models</i>	
Xiaoyan Zhao, Lingzhi Wang, Zhanghao Wang, Hong Cheng, Rui Zhang and Kam-Fai Wong	12564
<i>PAD: A Robustness Enhancement Ensemble Method via Promoting Attention Diversity</i>	
Yuting Yang, Pei Huang, Feifei Ma, Juan Cao and Jintao Li	12574
<i>Palmyra 3.0: A User-Friendly Cloud-Based Platform for Morphology and Dependency Syntax Annotation</i>	
Muhammed AbuOdeh, Long Phan, Ahmed Farouk Zakaria Elshabrawy and Nizar Habash	12585
<i>Parameter-Efficient Transfer Learning for End-to-end Speech Translation</i>	
Yunlong Zhao, Kexin Wang, Qianqian Dong and Tom Ko	12592
<i>ParaNames 1.0: Creating an Entity Name Corpus for 400+ Languages Using Wikidata</i>	
Jonne Sällevä and Constantine Lignos	12599
<i>PaReNT (Parent Retrieval Neural Tool): A Deep Dive into Word Formation across Languages</i>	
Emil Svoboda and Magda Sevcikova	12611
<i>Parsing for Mauritian Creole Using Universal Dependencies</i>	
Neha Ramsurrun, Rolando Coto-Solano and Michael Gonzalez	12622
<i>Parsing Headed Constituencies</i>	
Katarzyna Krasnowska-Kieraś and Marcin Woliński	12633
<i>PASUM: A Pre-training Architecture for Social Media User Modeling Based on Text Graph</i>	
Kun Wu, Xinyi Mou, Lanqing Xue, Zhenzhe Ying, Weiqiang Wang, Qi Zhang, Xuanjing Huang and Zhongyu Wei	12644
<i>Pater Incertus? There Is a Solution: Automatic Discrimination between Cognates and Borrowings for Romance Languages</i>	
Liviu P. Dinu, Ana Sabina Uban, Ioan-Bogdan Iordache, Alina Maria Cristea, Simona Georgescu and Laurentiu Zoicas	12657

<i>PDAMeta: Meta-Learning Framework with Progressive Data Augmentation for Few-Shot Text Classification</i>	
Xurui Li, Kaisong Song, Tianqianjing Lin, Yangyang Kang, Fubang Zhao, Changlong Sun and Xiaozhong Liu	12668
<i>PEaCE: A Chemistry-Oriented Dataset for Optical Character Recognition on Scientific Documents</i>	
Nan Zhang, Connor Heaton, Sean Timothy Okonsky, Prasenjit Mitra and Hilal Ezgi Toraman	12679
<i>PECC: Problem Extraction and Coding Challenges</i>	
Patrick Haller, Jonas Golde and Alan Akbik	12690
<i>PejorativITy: Disambiguating Pejorative Epithets to Improve Misogyny Detection in Italian Tweets</i>	
Arianna Muti, Federico Ruggeri, Cagri Toraman, Alberto Barrón-Cedeño, Samuel Algherini, Lorenzo Musetti, Silvia Ronchi, Gianmarco Saretto and Caterina Zapparoli	12700
<i>Persona-aware Multi-party Conversation Response Generation</i>	
Khyati Mahajan and Samira Shaikh	12712
<i>Phonetic Segmentation of the UCLA Phonetics Lab Archive</i>	
Eleanor Chodroff, Blaž Pažon, Annie Baker and Steven Moran	12724
<i>Phonotactic Complexity across Dialects</i>	
Ryan Soh-Eun Shim, Calvin Chang and David R. Mortensen	12734
<i>PILA: A Historical-Linguistic Dataset of Proto-Italic and Latin</i>	
Stephen Bothwell, Brian DuSell, David Chiang and Brian Krostenko	12749
<i>PIRB: A Comprehensive Benchmark of Polish Dense and Hybrid Text Retrieval Methods</i>	
Slawomir Dadas, Michał Peretkiewicz and Rafał Poświata	12761
<i>PLAES: Prompt-generalized and Level-aware Learning Framework for Cross-prompt Automated Essay Scoring</i>	
Yuan Chen and Xia Li	12775
<i>Plots Made Quickly: An Efficient Approach for Generating Visualizations from Natural Language Queries</i>	
Henrik Voigt, Kai Lawonn and Sina Zarriß	12787
<i>Pluggable Neural Machine Translation Models via Memory-augmented Adapters</i>	
Yuzhuang Xu, Shuo Wang, Peng Li, Xuebo Liu, Xiaolong Wang, Weidong Liu and Yang Liu	12794
<i>Pointing Out the Shortcomings of Relation Extraction Models with Semantically Motivated Adversarials</i>	
Gennaro Nolano, Moritz Blum, Basil Ell and Philipp Cimiano	12809
<i>Polish-ASTE: Aspect-Sentiment Triplet Extraction Datasets for Polish</i>	
Marta Lango, Borys Naglik, Mateusz Lango and Iwo Naglik	12821

<i>Polish Discourse Corpus (PDC): Corpus Design, ISO-Compliant Annotation, Data Highlights, and Parser Development</i>	
Maciej Ogrodniczuk, Aleksandra Tomaszewska, Daniel Ziembicki, Sebastian Żurowski, Ryszard Tuora and Aleksandra Zwierzchowska	12829
<i>PolitiCause: An Annotation Scheme and Corpus for Causality in Political Texts</i>	
Paulina Garcia Corral, Hanna Bechara, Ran Zhang and Slava Jankin	12836
<i>PolQA: Polish Question Answering Dataset</i>	
Piotr Rybak, Piotr Przybyła and Maciej Ogrodniczuk	12846
<i>PolyNERE: A Novel Ontology and Corpus for Named Entity Recognition and Relation Extraction in Polymer Science Domain</i>	
Van-Thuy Phi, Hiroki Teranishi, Yuji Matsumoto, Hiroyuki Oka and Masashi Ishii . . .	12856
<i>PopALM: Popularity-Aligned Language Models for Social Media Trendy Response Prediction</i>	
Erxin Yu, Jing Li and Chunpu Xu	12867
<i>PopAut: An Annotated Corpus for Populism Detection in Austrian News Comments</i>	
Ahmadou Wagne, Julia Neidhardt and Thomas Elmar Kolb	12879
<i>Positive and Risky Message Assessment for Music Products</i>	
Yigeng Zhang, Mahsa Shafaei, Fabio Gonzalez and Tamar Solorio	12893
<i>POS Tagging for the Endangered Dagur Language</i>	
Joanna Dolińska and Delphine Bernhard	12906
<i>Post-decoder Biasing for End-to-End Speech Recognition of Multi-turn Medical Interview</i>	
Heyang Liu, Yanfeng Wang and Yu Wang	12917
<i>PPORTAL_ner: An Annotated Corpus of Portuguese Literary Entities</i>	
Mariana O. Silva and Mirella M. Moro	12927
<i>Predictive and Distinctive Linguistic Features in Schizophrenia-Bipolar Spectrum Disorders</i>	
Martina Katalin Szabó, Veronika Vincze, Bernadett Dam, Csenge Guba, Anita Bagi and István Szendi	12938
<i>Prefix-diffusion: A Lightweight Diffusion Model for Diverse Image Captioning</i>	
Guisheng Liu, Yi Li, Zhengcong Fei, Haiyan Fu, Xiangyang Luo and Yanqing Guo .	12954
<i>Pre-Trained Language Models Represent Some Geographic Populations Better than Others</i>	
Jonathan Dunn, Benjamin Adams and Harish Tayyar Madabushi	12966
<i>Pre-training Cross-Modal Retrieval by Expansive Lexicon-Patch Alignment</i>	
Yang Yiyuan, Guodong Long, Michael Blumenstein, Xiubo Geng, Chongyang Tao, Tao Shen and Daxin Jiang	12977
<i>PRIMO: Progressive Induction for Multi-hop Open Rule Generation</i>	
Jianyu Liu, Sheng Bi and Guilin Qi	12988
<i>Principal Component Analysis as a Sanity Check for Bayesian Phylolinguistic Reconstruction</i>	
Yugo Murawaki	12999

<i>Prior Relational Schema Assists Effective Contrastive Learning for Inductive Knowledge Graph Completion</i>	
Ruilin Luo, Jiayi Li, Jianghangfan Zhang, Jing Xiao and Yujiu Yang	13014
<i>Probe Then Retrieve and Reason: Distilling Probing and Reasoning Capabilities into Smaller Language Models</i>	
Yichun Zhao, Shuheng Zhou and Huijia Zhu	13026
<i>Probing Large Language Models for Scalar Adjective Lexical Semantics and Scalar Diversity Pragmatics</i>	
Fangru Lin, Daniel Altshuler and Janet B. Pierrehumbert	13033
<i>Probing Multimodal Large Language Models for Global and Local Semantic Representations</i>	
Mingxu Tao, Quzhe Huang, Kun Xu, Liwei Chen, Yansong Feng and Dongyan Zhao	13050
<i>ProCQA: A Large-scale Community-based Programming Question Answering Dataset for Code Search</i>	
Zehan Li, Jianfei Zhang, Chuantao Yin, Yuanxin Ouyang and Wenge Rong	13057
<i>PRODIS - a Speech Database and a Phoneme-based Language Model for the Study of Predictability Effects in Polish</i>	
Zofia Malisz, Jan Foremski and Małgorzata Kul	13068
<i>Producing a Parallel Universal Dependencies Treebank of Ancient Hebrew and Ancient Greek via Cross-Lingual Projection</i>	
Daniel G. Swanson, Bryce D. Bussert and Francis Tyers	13074
<i>Projective Methods for Mitigating Gender Bias in Pre-trained Language Models</i>	
Hillary Dawkins, Isar Nejadgholi, Daniel Gillis and Judi McCuaig	13079
<i>Project MOSLA: Recording Every Moment of Second Language Acquisition</i>	
Masato Hagiwara and Joshua B. Tanner	13092
<i>PROM: A Phrase-level Copying Mechanism with Pre-training for Abstractive Summarization</i>	
Xinbei Ma, Yeyun Gong, Pengcheng He, Hai Zhao and Nan Duan	13103
<i>PromISE: Releasing the Capabilities of LLMs with Prompt Introspective Search</i>	
Minzheng Wang, Nan Xu, Jiahao Zhao, Yin Luo and Wenji Mao	13120
<i>Prompt-based Generation of Natural Language Explanations of Synthetic Lethality for Cancer Drug Discovery</i>	
Ke Zhang, Yimiao Feng and Jie Zheng	13131
<i>Prompt-based Zero-shot Relation Extraction with Semantic Knowledge Augmentation</i>	
Jiaying Gong and Hoda Eldardiry	13143
<i>Prompt-fused Framework for Inductive Logical Query Answering</i>	
Zezhong Xu, Wen Zhang, Peng Ye, Lei Liang and Huajun Chen	13157
<i>Prompting-based Synthetic Data Generation for Few-Shot Question Answering</i>	
Maximilian Schmidt, Andrea Bartezzaghi and Ngoc Thang Vu	13168

<i>Prompting Explicit and Implicit Knowledge for Multi-hop Question Answering Based on Human Reading Process</i>	
Guangming Huang, Yunfei Long, Cunjin Luo, Jiaying Shen and Xia Sun	13179
<i>Prompting for Numerical Sequences: A Case Study on Market Comment Generation</i>	
Masayuki Kawarada, Tatsuya Ishigaki and Hiroya Takamura	13190
<i>Prompting Large Language Models for Counterfactual Generation: An Empirical Study</i>	
Yongqi Li, Mayi Xu, Xin Miao, Shen Zhou and Tiejun Qian.....	13201
<i>PromptStream: Self-Supervised News Story Discovery Using Topic-Aware Article Representations</i>	
Arezoo Hatefi, Anton Eklund and Mona Forsman	13222
<i>Prompt Tuning for Few-shot Relation Extraction via Modeling Global and Local Graphs</i>	
Zirui Zhang, Yiyu Yang and Benhui Chen.....	13233
<i>PrOnto: Language Model Evaluations for 859 Languages</i>	
Luke Gessler.....	13243
<i>Prophecy Distillation for Boosting Abstractive Summarization</i>	
Jiaxin Duan, Fengyu Lu and Junfei Liu	13257
<i>Prototype-based Prompt-Instance Interaction with Causal Intervention for Few-shot Event Detection</i>	
Jingyao Tang, Lishuang Li, Hongbin Lu, Xueyang Qin, Beibei Zhang and Haiming Wu.....	13269
<i>Pruning before Fine-tuning: A Retraining-free Compression Framework for Pre-trained Language Models</i>	
Pingjie Wang, Hongcheng Liu, Yanfeng Wang and Yu Wang	13279
<i>PSentScore: Evaluating Sentiment Polarity in Dialogue Summarization</i>	
Yongxin Zhou, Fabien Ringeval and François Portet	13290
<i>Pseudonymization Categories across Domain Boundaries</i>	
Maria Irena Szawerna, Simon Dobnik, Therese Lindström Tiedemann, Ricardo Muñoz Sánchez, Xuan-Son Vu and Elena Volodina	13303
<i>PSE v1.0: The First Open Access Corpus of Public Service Encounters</i>	
Ingrid Espinoza, Steffen Frenzel, Laurin Friedrich, Wassiliki Siskou, Steffen Eckhard and Annette Hautli-Janisz	13315
<i>PSYDIAL: Personality-based Synthetic Dialogue Generation Using Large Language Models</i>	
Ji-Eun Han, Jun-Seok Koh, Hyeon-Tae Seo, Du-Seong Chang and Kyung-Ah Sohn.....	13321
<i>Puntuguese: A Corpus of Puns in Portuguese with Micro-edits</i>	
Marcio Lima Inacio, Gabriela Wick-Pedro, Renata Ramisch, Luís Espírito Santo, Xiomara S. Q. Chacon, Roney Santos, Rogério Sousa, Rafael Anchiêta and Hugo Goncalo Oliveira.....	13332
<i>PWESuite: Phonetic Word Embeddings and Tasks They Facilitate</i>	
Vilém Zouhar, Calvin Chang, Chenxuan Cui, Nate B. Carlson, Nathaniel Romney Robinson, Mrinmaya Sachan and David R. Mortensen	13344

<i>PyRater: A Python Toolkit for Annotation Analysis</i> Angelo Basile, Marc Franco-Salvador and Paolo Rosso.....	13356
<i>Qabas: An Open-Source Arabic Lexicographic Database</i> Mustafa Jarrar and Tymaa Hasanain Hammouda.....	13363
<i>QA-based Event Start-Points Ordering for Clinical Temporal Relation Annotation</i> Seiji Shimizu, Lis Pereira, Shuntaro Yada and Eiji Aramaki.....	13371
<i>QCAW 1.0: Building a Qatari Corpus of Student Argumentative Writing</i> Wajdi Zaghouni, Abdelhamid Ahmed, Xiao Zhang and Lameya Rezk.....	13382
<i>QDMR-based Planning-and-Solving Prompting for Complex Reasoning Tasks</i> Jinfeng Huang, Qiaoqiao She, Wenbin Jiang, Hua Wu, Yang Hao, Tong Xu and Feng Wu	13395
<i>Qsnail: A Questionnaire Dataset for Sequential Question Generation</i> Yan Lei, Liang Pang, Yuanzhuo Wang, Huawei Shen and Xueqi Cheng.....	13407
<i>Quantifying the Impact of Disfluency on Spoken Content Summarization</i> Maria Teleki, Xiangjue Dong and James Caverlee.....	13419
<i>QUEEREOTYPES: A Multi-Source Italian Corpus of Stereotypes towards LGBTQIA+ Community Members</i> Alessandra Teresa Cignarella, Manuela Sanguinetti, Simona Frenda, Andrea Marra, Cristina Bosco and Valerio Basile.....	13429
<i>Query-driven Relevant Paragraph Extraction from Legal Judgments</i> Santosh T.Y.S.S., Elvin A. Quero Hernandez and Matthias Grabmair.....	13442
<i>QueryNER: Segmentation of E-commerce Queries</i> Chester Palen-Michel, Lizzie Liang, Zhe Wu and Constantine Lignos.....	13455
<i>Question Answering over Tabular Data with DataBench: A Large-Scale Empirical Evaluation of LLMs</i> Jorge Osés Grijalba, L. Alfonso Ureña-López, Eugenio Martínez Cámara and Jose Camacho-Collados.....	13471
<i>Quite Good, but Not Enough: Nationality Bias in Large Language Models - a Case Study of ChatGPT</i> Shucheng Zhu, Weikang Wang and Ying Liu.....	13489
<i>RAAMove: A Corpus for Analyzing Moves in Research Article Abstracts</i> Hongzheng Li, Ruojin Wang, Ge Shi, Xing Lv, Lei Lei, Chong Feng, Fang Liu, Jinkun Lin, Yangguang Mei and Linnan Xu.....	13503
<i>RADCoT: Retrieval-Augmented Distillation to Specialization Models for Generating Chain-of-Thoughts in Query Expansion</i> Sung-Min Lee, Eunhwan Park, DongHyeon Jeon, Inho Kang and Seung-Hoon Na.....	13514
<i>RankPrompt: Step-by-Step Comparisons Make Language Models Better Reasoners</i> Chi Hu, Yuan Ge, Xiangnan Ma, Hang Cao, Qiang Li, Yonghua Yang, Tong Xiao and Jingbo Zhu.....	13524

<i>Rapidly Developing High-quality Instruction Data and Evaluation Benchmark for Large Language Models with Minimal Human Effort: A Case Study on Japanese</i> Yikun Sun, Zhen Wan, Nobuhiro Ueda, Sakiko Yahata, Fei Cheng, Chenhui Chu and Sadao Kurohashi	13537
<i>Rapidly Piloting Real-time Linguistic Assistance for Simultaneous Interpreters with Untrained Bilingual Surrogates</i> Alvin C. Grissom II, Jo Shoemaker, Benjamin Goldman, Ruikang Shi, Craig Stewart, C. Anton Rytting, Leah Findlater and Jordan Boyd-Graber	13548
<i>Rationale-based Learning Using Self-Supervised Narrative Events for Text Summarisation of Interactive Digital Narratives</i> Ashwathy T Revi, Stuart E. Middleton and David E. Millard	13557
<i>Reading Does Not Equal Reading: Comparing, Simulating and Exploiting Reading Behavior across Populations</i> David R. Reich, Shuwen Deng, Marina Björnsdóttir, Lena Jäger and Nora Hollenstein	13586
<i>ReadLet: A Dataset for Oral, Visual and Tactile Text Reading Data of Early and Mature Readers</i> Marcello Ferro, Claudia Marzi, Andrea Nadalini, Loukia Taxitari, Alessandro Lento and Vito Pirrelli	13595
<i>Reassessing Semantic Knowledge Encoded in Large Language Models through the Word-in-Context Task</i> Yoshihiko Hayashi	13610
<i>Rebalancing Label Distribution While Eliminating Inherent Waiting Time in Multi Label Active Learning Applied to Transformers</i> Maxime Arens, Lucile Callebert, Mohand Boughanem and Jose G. Moreno	13621
<i>ReCAP: Semantic Role Enhanced Caption Generation</i> Abhidip Bhattacharyya, Martha Palmer and Christoffer Heckman	13633
<i>Recent Trends in Personalized Dialogue Generation: A Review of Datasets, Methodologies, and Evaluations</i> Yi-Pei Chen, Noriki Nishida, Hideki Nakayama and Yuji Matsumoto	13650
<i>RECIPE4U: Student-ChatGPT Interaction Dataset in EFL Writing Education</i> Jieun Han, Haneul Yoo, Junho Myung, Minsun Kim, Tak Yeon Lee, So-Yeon Ahn and Alice Oh	13666
<i>Recognizing Social Cues in Crisis Situations</i> Di Wang, Yuan Zhuang, Ellen Riloff and Marina Kogan	13677
<i>Recognizing Value Resonance with Resonance-Tuned RoBERTa Task Definition, Experimental Validation, and Robust Modeling</i> Noam K. Benkler, Scott Friedman, Sonja Schmer-Galunder, Drisana Marissa Mosaphir, Robert P. Goldman, Ruta Wheelock, Vasanth Sarathy, Pavan Kantharaju and Matthew D. McLure	13688
<i>Recommending Missed Citations Identified by Reviewers: A New Task, Dataset and Baselines</i> Kehan Long, Shasha Li, Pancheng Wang, Chenlong Bao, Jintao Tang and Ting Wang	13699

<i>Reconstruction of Cuneiform Literary Texts as Text Matching</i> Fabian Simonjetz, Jussi Laasonen, Yunus Cobanoglu, Alexander Fraser and Enrique Jiménez	13712
<i>Reduce Redundancy Then Rerank: Enhancing Code Summarization with a Novel Pipeline Framework</i> Xiaoyu Hu, Xu Zhang, Zexu Lin and Deyu Zhou	13722
<i>Re-evaluating the Tomes for the Times</i> Ryan Brate, Marieke van Erp and Antal van den Bosch	13734
<i>REFeREE: A REFERENCE-FREE Model-Based Metric for Text Simplification</i> Yichen Huang and Ekaterina Kochmar	13740
<i>Reference-guided Style-Consistent Content Transfer</i> Wei-Fan Chen, Milad Alshomary, Maja Stahl, Khalid Al Khatib, Benno Stein and Henning Wachsmuth	13754
<i>Reference-less Analysis of Context Specificity in Translation with Personalised Language Models</i> Sebastian Vincent, Rowanne Sumner, Alice Dowek, Charlotte Prescott, Emily Preston, Chris Bayliss, Chris Oakley and Carolina Scarton	13769
<i>Refining Idioms Semantics Comprehension via Contrastive Learning and Cross-Attention</i> Mingmin Wu, Guixin Su, Yongcheng Zhang, Zhongqiang Huang and Ying Sha	13785
<i>Refining rtMRI Landmark-Based Vocal Tract Contour Labels with FCN-Based Smoothing and Point-to-Curve Projection</i> Mushaffa Rasyid Ridha and Sakriani Sakti	13796
<i>Reflecting the Male Gaze: Quantifying Female Objectification in 19th and 20th Century Novels</i> Kexin Luo, Yue Mao, Bei Zhang and Sophie Hao	13803
<i>Reflections & Resonance: Two-Agent Partnership for Advancing LLM-based Story Annotation</i> Yuetian Chen and Mei Si	13813
<i>ReflectSumm: A Benchmark for Course Reflection Summarization</i> Mohamed Elaraby, Yang Zhong, Diane Litman, Ahmed Ashraf Butt and Muhsin Menekse	13819
<i>Reimagining Intent Prediction: Insights from Graph-Based Dialogue Modeling and Sentence Encoders</i> Daria Romanovna Ledneva and Denis Pavlovich Kuznetsov	13847
<i>Reinforcement Retrieval Leveraging Fine-grained Feedback for Fact Checking News Claims with Black-Box LLM</i> Xuan Zhang and Wei Gao	13861
<i>Related Work Is All You Need</i> Rodolfo Joel Zevallos, John E. Ortega and Benjamin Irving	13874
<i>Relation between Cross-Genre and Cross-Topic Transfer in Dependency Parsing</i> Vera Danilova and Sara Stymne	13879

<i>Relation Classification via Bidirectional Prompt Learning with Data Augmentation by Large Language Model</i>	
Yizhi Jiang, Jinlong Li and Huanhuan Chen	13885
<i>Release of Pre-Trained Models for the Japanese Language</i>	
Kei Sawada, Tianyu Zhao, Makoto Shing, Kentaro Mitsui, Akio Kaga, Yukiya Hono, Toshiaki Wakatsuki and Koh Mitsuda	13898
<i>Releasing the Capacity of GANs in Non-Autoregressive Image Captioning</i>	
Da Ren and Qing Li	13906
<i>RENN: A Rule Embedding Enhanced Neural Network Framework for Temporal Knowledge Graph Completion</i>	
Linlin Zong, Zhenrong Xie, Chi Ma, Xinyue Liu, Xianchao Zhang and Bo Xu	13919
<i>Replace, Paraphrase or Fine-tune? Evaluating Automatic Simplification for Medical Texts in Spanish</i>	
Leonardo Campillos-Llanos, Ana Rosa Terroba, Rocío Bartolomé, Ana Valverde-Mateos, Cristina González, Adrián Capllonch-Carrión and Jonathan Heras	13929
<i>Representation Degeneration Problem in Prompt-based Models for Natural Language Understanding</i>	
Qingyan Zhao, Ruifang He, Jinpeng Zhang, Chang Liu and Bo Wang	13946
<i>Representing Compounding with OntoLex. An Evaluation of Vocabularies for Word Formation Resources</i>	
Elena Benzoni, Matteo Pellegrini, Francesco Dedè and Marco Passarotti	13958
<i>Reranking Overgenerated Responses for End-to-End Task-Oriented Dialogue Systems</i>	
Songbo Hu, Ivan Vulić, Fangyu Liu and Anna Korhonen	13970
<i>Resolving Legalese: A Multilingual Exploration of Negation Scope Resolution in Legal Documents</i>	
Ramona Christen, Anastassia Shaitarova, Matthias Stürmer and Joel Niklaus	13992
<i>Restoring Ancient Ideograph: A Multimodal Multitask Neural Network Approach</i>	
Siyu Duan, Jun Wang and Qi Su	14005
<i>Retentive or Forgetful? Diving into the Knowledge Memorizing Mechanism of Language Models</i>	
Boxi Cao, Qiaoyu Tang, Hongyu Lin, Shanshan Jiang, Bin Dong, Xianpei Han, Jiawei Chen, Tianshu Wang and Le Sun	14016
<i>Rethinking Word-level Adversarial Attack: The Trade-off between Efficiency, Effectiveness, and Imperceptibility</i>	
Pengwei Zhan, Jing Yang, He Wang, Chao Zheng and Liming Wang	14037
<i>Retrieval-Augmented Modular Prompt Tuning for Low-Resource Data-to-Text Generation</i>	
Ruitao Feng, Xudong Hong, Mayank Jobanputra, Mattes Warning and Vera Demberg	14053
<i>Retrieval-based Question Answering with Passage Expansion Using a Knowledge Graph</i>	
Benno Kruit, Yiming Xu and Jan-Christoph Kalo	14063
<i>Revisiting Context Choices for Context-aware Machine Translation</i>	
Matiss Rikters and Toshiaki Nakazawa	14073

<i>Revisiting Data Reconstruction Attacks on Real-world Dataset for Federated Natural Language Understanding</i>	
Zhuo Zhang, Jintao Huang, Xiangjing Hu, Jingyuan Zhang, Yating Zhang, Hui Wang, Yue Yu, Qifan Wang, Lizhen Qu and Zenglin Xu	14080
<i>Revisiting the Classics: A Study on Identifying and Rectifying Gender Stereotypes in Rhymes and Poems</i>	
Aditya Narayan Sankaran, Vigneshwaran Shankaran, Sampath Lonka and Rajesh Sharma	14092
<i>Revisiting the Self-Consistency Challenges in Multi-Choice Question Formats for Large Language Model Evaluation</i>	
Wenjie Zhou, Qiang Wang, Mingzhou Xu, Ming Chen and Xiangyu Duan	14103
<i>Revisiting Three Text-to-Speech Synthesis Experiments with a Web-Based Audience Response System</i>	
Christina Tännander, Jens Edlund and Joakim Gustafson	14111
<i>Rewiring the Transformer with Depth-Wise LSTMs</i>	
Hongfei Xu, Yang Song, Qihui Liu, Josef van Genabith and Deyi Xiong	14122
<i>RISE: Robust Early-exiting Internal Classifiers for Suicide Risk Evaluation</i>	
Ritesh Singh Soun, Atula Tejaswi Neerkaje, Ramit Sawhney, Nikolaos Aletras and Preslav Nakov	14134
<i>RoBERTa Low Resource Fine Tuning for Sentiment Analysis in Albanian</i>	
Krenare Pireva Nuci, Paul Landes and Barbara Di Eugenio	14146
<i>RoboVox: A Single/Multi-channel Far-field Speaker Recognition Benchmark for a Mobile Robot</i>	
Mohammad Mohammadamini, Driss Matrouf, Michael Rouvier, Jean-Francois Bonastre, Romain Serizel and Theophile Gonos	14152
<i>Robust and Scalable Model Editing for Large Language Models</i>	
Yingfa Chen, Zhengyan Zhang, Xu Han, Chaojun Xiao, Zhiyuan Liu, Chen Chen, Kuai Li, Tao Yang and Maosong Sun	14157
<i>RoCode: A Dataset for Measuring Code Intelligence from Problem Definitions in Romanian</i>	
Adrian Cosma, Ioan-Bogdan Iordache and Paolo Rosso	14173
<i>RoColns: Enhancing Robustness of Large Language Models through Code-Style Instructions</i>	
Yuansen Zhang, Xiao Wang, Zhiheng Xi, Han Xia, Tao Gui, Qi Zhang and Xuanjing Huang	14186
<i>RT-VQ2A2: Real Time Vector Quantized Question Answering with ASR</i>	
Kyungho Kim, Seongmin Park and Jihwa Lee	14204
<i>RU22Fact: Optimizing Evidence for Multilingual Explainable Fact-Checking on Russia-Ukraine Conflict</i>	
Yirong Zeng, Xiao Ding, Yi Zhao, Xiangyu Li, Jie Zhang, Chao Yao, Ting Liu and Bing Qin	14215
<i>RuBia: A Russian Language Bias Detection Dataset</i>	
Veronika Grigoreva, Anastasiia Ivanova, Ilseyar Alimova and Ekaterina Artemova ..	14227

<i>Russian Learner Corpus: Towards Error-Cause Annotation for L2 Russian</i> Daniil Kosakin, Sergei Obiedkov, Ivan Smirnov, Ekaterina Rakhilina, Anastasia Vyrenkova and Ekaterina Zalivina	14240
<i>S3Prompt: Instructing the Model with Self-calibration, Self-recall and Self-aggregation to Improve In-context Learning</i> Junda Chen and Jianting Liu	14259
<i>SaGE: Evaluating Moral Consistency in Large Language Models</i> Vamshi Krishna Bonagiri, Sreeram Vennam, Priyanshu Govil, Ponnuram Kumaraguru and Manas Gaur	14272
<i>Saliency-Aware Interpolative Augmentation for Multimodal Financial Prediction</i> Samyak Jain, Parth Chhabra, Atula Tejaswi Neerkaje, Puneet Mathur, Ramit Sawhney, Shivam Agarwal, Preslav Nakov, Sudheer Chava and Dinesh Manocha	14285
<i>Samayik: A Benchmark and Dataset for English-Sanskrit Translation</i> Ayush Maheshwari, Ashim Gupta, Amrith Krishna, Atul Kumar Singh, Ganesh Ramakrishnan, Anil Kumar Gourishetty and Jitin Singla	14298
<i>SamróMur Milljón: An ASR Corpus of One Million Verified Read Prompts in Icelandic</i> Carlos Daniel Hernandez Mena, Þorsteinn Daði Gunnarsson and Jon Gudnason ..	14305
<i>Sarcasm Detection in a Disaster Context</i> Tiberiu Sosea, Junyi Jessy Li and Cornelia Caragea	14313
<i>SarcNet: A Multilingual Multimodal Sarcasm Detection Dataset</i> Tan Yue, Xuzhao Shi, Rui Mao, Zonghai Hu and Erik Cambria	14325
<i>Scalable Patent Classification with Aggregated Multi-View Ranking</i> Dan Li, Vikrant Yadav, Zi Long Zhu, Maziar Moradi Fard, Zubair Afzal and George Tsatsaronis	14336
<i>Scale-VAE: Preventing Posterior Collapse in Variational Autoencoder</i> Tianbao Song, Jingbo Sun, Xin Liu and Weiming Peng	14347
<i>Scaling Data Diversity for Fine-Tuning Language Models in Human Alignment</i> Feifan Song, Bowen Yu, Hao Lang, Haiyang Yu, Fei Huang, Houfeng Wang and Yongbin Li	14358
<i>Scansion-based Lyrics Generation</i> Yiwen Chen and Simone Teufel	14370
<i>Schema-based Data Augmentation for Event Extraction</i> Xiaomeng Jin and Heng Ji	14382
<i>Schema Learning Corpus: Data and Annotation Focused on Complex Events</i> Song Chen, Jennifer Tracey, Ann Bies and Stephanie Strassel	14393
<i>Schrodinger's Threshold: When the AUC Doesn't Predict Accuracy</i> Juri Opitz	14400
<i>SciDMT: A Large-Scale Corpus for Detecting Scientific Mentions</i> Huitong Pan, Qi Zhang, Cornelia Caragea, Eduard Dragut and Longin Jan Latecki .	14407

<i>SciMRC: Multi-perspective Scientific Machine Reading Comprehension</i> Xiao Zhang, Heqi Zheng, Yuxiang Nie, Heyan Huang and Xian-Ling Mao	14418
<i>SciNews: From Scholarly Complexities to Public Narratives – a Dataset for Scientific News Report Generation</i> Dongqi Pu, Yifan Wang, Jia E. Loy and Vera Demberg.....	14429
<i>SCOUT: A Situated and Multi-Modal Human-Robot Dialogue Corpus</i> Stephanie M. Lukin, Claire Bonial, Matthew Marge, Taylor A. Hudson, Cory J. Hayes, Kimberly Pollard, Anthony Baker, Ashley N. Fouts, Ron Artstein, Felix Gervits, Mitchell Abrams, Cassidy Henry, Lucia Donatelli, Anton Leuski, Susan G. Hill, David Traum and Clare Voss	14445
<i>SDA: Simple Discrete Augmentation for Contrastive Sentence Representation Learning</i> Dongsheng Zhu, Zhenyu Mao, Jinghui Lu, Rui Zhao and Fei Tan	14459
<i>Searching by Code: A New SearchBySnippet Dataset and SnippeR Retrieval Model for Searching by Code Snippets</i> Ivan Sedykh, Nikita Sorokin, Dmitry Abulkhanov, Sergey I. Nikolenko and Valentin Malykh	14472
<i>Sebastian, Basti, Wast!?! Recognizing Named Entities in Bavarian Dialectal Data</i> Siyao Peng, Zihang Sun, Huangyan Shan, Marie Kolm, Verena Blaschke, Ekaterina Artemova and Barbara Plank	14478
<i>Seeing Eye-to-Eye: Cross-Modal Coherence Relations Inform Eye-gaze Patterns During Comprehension & Production</i> Mert Inan and Malihe Alikhani	14494
<i>Seeing Is Believing! towards Knowledge-Infused Multi-modal Medical Dialogue Generation</i> Abhisek Tiwari, Shreyangshu Bera, Preeti Verma, Jaithra Varma Manthena, Sriparna Saha, Pushpak Bhattacharyya, Minakshi Dhar and Sarbajeet Tiwari	14513
<i>Segmentation of Complex Question Turns for Argument Mining: A Corpus-based Study in the Financial Domain</i> Giulia D’Agostino, Chris A. Reed and Daniele Puccinelli	14524
<i>Select and Reorder: A Novel Approach for Neural Sign Language Production</i> Harry Walsh, Ben Saunders and Richard Bowden	14531
<i>Select High-quality Synthetic QA Pairs to Augment Training Data in MRC under the Reward Guidance of Generative Language Models</i> Jing Jin and Houfeng Wang	14543
<i>Selective Temporal Knowledge Graph Reasoning</i> Zhongni Hou, Xiaolong Jin, Zixuan Li, Long Bai, Jiafeng Guo and Xueqi Cheng	14555
<i>Self-Explanation Prompting Improves Dialogue Understanding in Large Language Models</i> Haoyu Gao, Ting-En Lin, Hangyu Li, Min Yang, Yuchuan Wu, Wentao Ma, Fei Huang and Yongbin Li	14567
<i>Self-Improvement Programming for Temporal Knowledge Graph Question Answering</i> Zhuo Chen, Zhao Zhang, Zixuan Li, Fei Wang, Yutao Zeng, Xiaolong Jin and Yongjun Xu	14579

<i>Self-Knowledge Distillation for Knowledge Graph Embedding</i> Haotian Xu, Yuhua Wang and Jiahui Fan	14595
<i>Self-reported Demographics and Discourse Dynamics in a Persuasive Online Forum</i> Agnieszka Falenska, Eva Maria Vecchi and Gabriella Lapesa.....	14606
<i>Semantic Frame Extraction in Multilingual Olfactory Events</i> Stefano Menini	14622
<i>Semantic Map-based Generation of Navigation Instructions</i> Chengzu Li, Chao Zhang, Simone Teufel, Rama Sanand Doddipatla and Svetlana Stoyanchev	14628
<i>Semantic Role Labeling Guided Out-of-distribution Detection</i> Jinan Zou, Maihao Guo, Yu Tian, Yuhao Lin, Haiyao Cao, Lingqiao Liu, Ehsan Abbasnejad and Javen Qinfeng Shi.....	14641
<i>Semantics-Aware Dual Graph Convolutional Networks for Argument Pair Extraction</i> Minzhao Guan, Zhixun Qiu, Fenghuan Li and Yun Xue	14652
<i>Semantics-enhanced Cross-modal Masked Image Modeling for Vision-Language Pre-training</i> Haowei Liu, Yaya Shi, Haiyang Xu, Chunfeng Yuan, Qinghao Ye, Chenliang Li, Ming Yan, Ji Zhang, Fei Huang, Bing Li and Weiming Hu.....	14664
<i>Sense of the Day: Short Timeframe Temporal-Aware Word Sense Disambiguation</i> Yuchen Wei and Milton King.....	14676
<i>SENTA: Sentence Simplification System for Slovene</i> Aleš Žagar, Matej Klemen, Marko Robnik-Šikonja and Iztok Kosem	14687
<i>SentiCSE: A Sentiment-aware Contrastive Sentence Embedding Framework with Sentiment-guided Textual Similarity</i> Jaemin Kim, Yohan Na, Kangmin Kim, Sang-Rak Lee and Dong-Kyu Chae	14693
<i>Sequence Reducible Holdout Loss for Language Model Pretraining</i> Raghuveer Thirukovalluru, Nicholas Monath, Bhuwan Dhingra and Sam Wiseman .	14705
<i>Sequence-to-Sequence Language Models for Character and Emotion Detection in Dream Narratives</i> Gustave Cortal	14717
<i>Sequence-to-Sequence Spanish Pre-trained Language Models</i> Vladimir Araujo, Maria Mihaela Trusca, Rodrigo Tufiño and Marie-Francine Moens .	14729
<i>Sequential and Repetitive Pattern Learning for Temporal Knowledge Graph Reasoning</i> Xuefei Li, Huiwei Zhou, Weihong Yao, Wenchu Li, Yingyu Lin and Lei Du	14744
<i>SGCM: Saliency-Guided Context Modeling for Question Generation</i> Chuyao Ding, Yu Hong and Jianmin Yao	14755
<i>ShadowSense: A Multi-annotated Dataset for Evaluating Word Sense Induction</i> Ondřej Herman and Miloš Jakubiček	14763

<i>Sharing the Cost of Success: A Game for Evaluating and Learning Collaborative Multi-Agent Instruction Giving and Following Policies</i>	
Philipp Sadler, Sherzod Hakimov and David Schlangen	14770
<i>SIGA: A Naturalistic NLI Dataset of English Scalar Implicatures with Gradable Adjectives</i>	
Rashid Nizamani, Sebastian Schuster and Vera Demberg	14784
<i>SignBLEU: Automatic Evaluation of Multi-channel Sign Language Translation</i>	
Jung-Ho Kim, Mathew John Huerta-Enochian, Changyong Ko and Du Hui Lee	14796
<i>SilverAlign: MT-Based Silver Data Algorithm for Evaluating Word Alignment</i>	
Abdullatif Koksai, Silvia Severini and Hinrich Schütze	14812
<i>Silver Retriever: Advancing Neural Passage Retrieval for Polish Question Answering</i>	
Piotr Rybak and Maciej Ogrodniczuk	14826
<i>SimLex-999 for Dutch</i>	
Lizzy Brans and Jelke Bloem	14832
<i>Sinkhorn Distance Minimization for Knowledge Distillation</i>	
Xiao Cui, Yulei Qin, Yuting Gao, Enwei Zhang, Zihan Xu, Tong Wu, Ke Li, Xing Sun, Wengang Zhou and Houqiang Li	14846
<i>SI-NLI: A Slovene Natural Language Inference Dataset and Its Evaluation</i>	
Matej Klemen, Aleš Žagar, Jaka Čibej and Marko Robnik-Šikonja	14859
<i>SkOTaPA: A Dataset for Skepticism Detection in Online Text after Persuasion Attempt</i>	
Smitha Muthya Sudheendra, Maral Abdollahi, Dongyeop Kang, Jisu Huh and Jaideep Srivastava	14871
<i>SLaCAD: A Spoken Language Corpus for Early Alzheimer's Disease Detection</i>	
Shahla Farzana, Edoardo Stoppa, Alex Leow, Tamar Gollan, Raeanne Moore, David Salmon, Douglas Galasko, Erin Sundermann and Natalie Parde	14877
<i>Slot and Intent Detection Resources for Bavarian and Lithuanian: Assessing Translations vs Natural Queries to Digital Assistants</i>	
Miriam Winkler, Virginija Juozapaityte, Rob van der Goot and Barbara Plank	14898
<i>SlovakSum: A Large Scale Slovak Summarization Dataset</i>	
Viktoria Ondrejova and Marek Suppa	14916
<i>Small Language Models Are Good Too: An Empirical Study of Zero-Shot Classification</i>	
Pierre Lepagnol, Thomas Gerald, Sahar Ghannay, Christophe Servan and Sophie Rosset	14923
<i>SmartTrim: Adaptive Tokens and Attention Pruning for Efficient Vision-Language Models</i>	
Zekun Wang, Jingchang Chen, Wangchunshu Zhou, Haichao Zhu, Jiafeng Liang, Liping Shan, Ming Liu, Dongliang Xu, Qing Yang and Bing Qin	14937
<i>SM-FEEL-BG - the First Bulgarian Datasets and Classifiers for Detecting Feelings, Emotions, and Sentiments of Bulgarian Social Media Text</i>	
Irina Temnikova, Iva Marinova, Silvia Gargova, Ruslana Margova, Alexander Komarov, Tsvetelina Stefanova, Veneta Kireva, Dimana Vyatrova, Nevena Grigorova, Yordan Mandevski and Stefan Minkov	14954

<i>SOBR: A Corpus for Stylometry, Obfuscation, and Bias on Reddit</i> Chris Emmerly, Marilù Miotto, Sergey Kramp and Bennett Kleinberg.....	14967
<i>Social Convos: Capturing Agendas and Emotions on Social Media</i> Ankita Bhaumik, Ning Sa, Gregorios Katsios and Tomek Strzalkowski.....	14984
<i>Social Orientation: A New Feature for Dialogue Analysis</i> Todd Morrill, Zhaoyuan Deng, Yanda Chen, Amith Ananthram, Colin Wayne Leach and Kathleen McKeown.....	14995
<i>SoftMCL: Soft Momentum Contrastive Learning for Fine-grained Sentiment-aware Pre-training</i> Jin Wang, Liang-Chih Yu and Xuejie Zhang.....	15012
<i>Soft-Prompting with Graph-of-Thought for Multi-modal Representation Learning</i> Jun Cheng Yang, Zuchao Li, Shuai Xie, Wei Yu, Shijun Li and Bo Du.....	15024
<i>Soft Well-Formed Semantic Parsing with Score-Based Selection</i> Jiangming Liu.....	15037
<i>So Hateful! Building a Multi-Label Hate Speech Annotated Arabic Dataset</i> Wajdi Zaghouani, Hamdy Mubarak and Md. Rafiul Biswas.....	15044
<i>Sonos Voice Control Bias Assessment Dataset: A Methodology for Demographic Bias Assessment in Voice Assistants</i> Chloe Sekkat, Fanny Leroy, Salima Mdhaffar, Blake Perry Smith, Yannick Estève, Joseph Dureau and Alice Coucke.....	15056
<i>Source-free Domain Adaptation for Aspect-based Sentiment Analysis</i> Zishuo Zhao, Ziyang Ma, Zhenzhou Lin, Jingyou Xie, Yinghui Li and Ying Shen.....	15076
<i>SPACE-IDEAS: A Dataset for Salient Information Detection in Space Innovation</i> Andres Garcia-Silva, Cristian Berrio and Jose Manuel Gomez-Perez.....	15087
<i>Spanish Resource Grammar Version 2023</i> Olga Zamaraeva, Lorena S. Allegue and Carlos Gómez-Rodríguez.....	15093
<i>Spanless Event Annotation for Corpus-Wide Complex Event Understanding</i> Ann Bies, Jennifer Tracey, Ann O'Brien, Song Chen and Stephanie Strassel.....	15105
<i>Sparse Logistic Regression with High-order Features for Automatic Grammar Rule Extraction from Treebanks</i> Santiago Herrera, Caio Corro and Sylvain Kahane.....	15114
<i>Specifying Genericity through Inclusiveness and Abstractness Continuous Scales</i> Claudia Collacciani, Andrea Amelio Ravelli and Marianna Bolognesi.....	15126
<i>SpeechAlign: A Framework for Speech Translation Alignment Evaluation</i> Belen Alastruey, Aleix Sant, Gerard I. Gállego, David Dale and Marta R. Costa-jussà.....	15137
<i>Speech Analysis of Language Varieties in Italy</i> Moreno La Quatra, Alkis Koudounas, Elena Baralis and Sabato Marco Siniscalchi.....	15147

<i>Speech Corpus for Korean Children with Autism Spectrum Disorder: Towards Automatic Assessment Systems</i>	
Seonwoo Lee, Jihyun Mun, Sunhee Kim and Minhwa Chung	15160
<i>Speech Recognition Corpus of the Khinalug Language for Documenting Endangered Languages</i>	
Zhaolin Li, Monika Rind-Pawłowski and Jan Niehues	15171
<i>SPICED: News Similarity Detection Dataset with Multiple Topics and Complexity Levels</i>	
Elena Shushkevich, Long Thanh Mai, Manuel V. Loureiro, Steven Derby and Tri Kurniawan Wijaya	15181
<i>SPLICE: A Singleton-Enhanced PipeLine for Coreference REsolution</i>	
Yilun Zhu, Siyao Peng, Sameer Pradhan and Amir Zeldes	15191
<i>SPOTTER: A Framework for Investigating Convention Formation in a Visually Grounded Human-Robot Reference Task</i>	
Jaap Kruijt, Peggy van Minkelen, Lucia Donatelli, Piek T.J.M. Vossen, Elly Konijn and Thomas Baier	15202
<i>SpreadNaLa: A Naturalistic Code Generation Evaluation Dataset of Spreadsheet Formulas</i>	
Sebastian Schuster, Ayesha Ansar, Om Agarwal and Vera Demberg	15216
<i>STAF: Pushing the Boundaries of Test-Time Adaptation towards Practical Noise Scenarios</i>	
Haoyu Xiong, Xinchun Zhang, Leixin Yang, Yu Xiang and Gang Fang	15226
<i>STAGE: Simple Text Data Augmentation by Graph Exploration</i>	
Ho-Seung Kim, YongHoon Kang and Jee-Hyong Lee	15238
<i>Stance Reasoner: Zero-Shot Stance Detection on Social Media with Explicit Reasoning</i>	
Maksym Taranukhin, Vered Shwartz and Evangelos Milios	15257
<i>STEntConv: Predicting Disagreement between Reddit Users with Stance Detection and a Signed Graph Convolutional Network</i>	
Isabelle Lorge, Li Zhang, Xiaowen Dong and Janet Pierrehumbert	15273
<i>Step-by-Step: Controlling Arbitrary Style in Text with Large Language Models</i>	
Pusheng Liu, Lianwei Wu, Linyong Wang, Sensen Guo and Yang Liu	15285
<i>Step Feasibility-Aware and Error-Correctable Entailment Tree Generation</i>	
Junyue Song, Xin Wu and Yi Cai	15296
<i>Still All Greeklsh to Me: Greeklsh to Greek Transliteration</i>	
Anastasios Toumazatos, John Pavlopoulos, Ion Androutsopoulos and Stavros Vassos	15309
<i>Stories and Personal Experiences in the COVID-19 Discourse</i>	
Neele Falk and Gabriella Lapesa	15320
<i>Strengthening the WiC: New Polysemy Dataset in Hindi and Lack of Cross Lingual Transfer</i>	
Haim Dubossarsky and Farheen Dairkee	15341
<i>StructAM: Enhancing Address Matching through Semantic Understanding of Structure-aware Information</i>	
Zhaoqi Zhang, Pasquale Balsebre, Siqiang Luo, Zhen Hai and Jiangping Huang ...	15350

<i>Structure-aware Fine-tuning for Code Pre-trained Models</i> Jiayi Wu, Renyu Zhu, Nuo Chen, Qiushi Sun, Xiang Li and Ming Gao	15362
<i>Structure-aware Generation Model for Cross-Domain Aspect-based Sentiment Classification</i> Shichen Li, Zhongqing Wang, Yanzhi Xu and Guodong Zhou	15373
<i>StyleFlow: Disentangle Latent Representations via Normalizing Flow for Unsupervised Text Style Transfer</i> Kangchen Zhu, Zhiliang Tian, Jingyu Wei, Ruifeng Luo, Yiping Song and Xiaoguang Mao	15384
<i>Submodular-based In-context Example Selection for LLMs-based Machine Translation</i> Baijun Ji, Xiangyu Duan, Zhenyu Qiu, Tong Zhang, Junhui Li, Hao Yang and Min Zhang	15398
<i>Subspace Defense: Discarding Adversarial Perturbations by Learning a Subspace for Clean Signals</i> Rui Zheng, Yuhao Zhou, Zhiheng Xi, Tao Gui, Qi Zhang and Xuanjing Huang	15410
<i>Sub-Table Rescorer for Table Question Answering</i> Atsushi Kojima	15422
<i>SUK 1.0: A New Training Corpus for Linguistic Annotation of Modern Standard Slovene</i> Špela Arhar Holdt, Jaka Čibej, Kaja Dobrovoljc, Tomaž Erjavec, Polona Gantar, Simon Krek, Tina Munda, Nejc Robida, Luka Terčon and Slavko Zitnik	15428
<i>SuperST: Superficial Self-Training for Few-Shot Text Classification</i> Ju-Hyoung Lee, Joonghyuk Hahn, Hyeon-Tae Seo, Jiho Park and Yo-Sub Han	15436
<i>SwissSLI: The Multi-parallel Sign Language Corpus for Switzerland</i> Zifan Jiang, Anne Göhring, Amit Moryossef, Rico Sennrich and Sarah Ebling	15448
<i>Synergetic Interaction Network with Cross-task Attention for Joint Relational Triple Extraction</i> Da Luo, Run Lin, Qiao Liu, Yuxiang Cai, Xueyi Liu, Yanglei Gan and Rui Hou	15457
<i>SynPrompt: Syntax-aware Enhanced Prompt Engineering for Aspect-based Sentiment Analysis</i> Wen Yin, Cencen Liu, Yi Xu, Ahmad Raza Wahla, Huang Yiting and Dezhong Zheng	15469
<i>Synthetic Data Generation and Joint Learning for Robust Code-Mixed Translation</i> Kartik Kartik, Sanjana Soni, Anoop Kunchukuttan, Tanmoy Chakraborty and Md. Shad Akhtar	15480
<i>SynTOD: Augmented Response Synthesis for Robust End-to-End Task-Oriented Dialogue System</i> Nguyen Quang Chieu, Quang-Minh Tran and Khac-Hoai Nam Bui	15493
<i>Tackling Long Code Search with Splitting, Encoding, and Aggregating</i> Fan Hu, Yanlin Wang, Lun Du, Hongyu Zhang, Dongmei Zhang and Xirong Li	15500
<i>TacoERE: Cluster-aware Compression for Event Relation Extraction</i> Yong Guan, Xiaozhi Wang, Lei Hou, Juanzi Li, Jeff Z. Pan, Jiaoyan Chen and Freddy Lecue	15511

<i>TACO – Twitter Arguments from COversations</i>	
Marc Feger and Stefan Dietze	15522
<i>TAeKD: Teacher Assistant Enhanced Knowledge Distillation for Closed-Source Multilingual Neural Machine Translation</i>	
Bo Lv, Xin Liu, Kaiwen Wei, Ping Luo and Yue Yu	15530
<i>TaiChi: Improving the Robustness of NLP Models by Seeking Common Ground While Reserving Differences</i>	
Huimin Chen, Chengyu Wang, Yanhao Wang, Gen Chen and Yinggui Wang	15542
<i>Take Care of Your Prompt Bias! Investigating and Mitigating Prompt Bias in Factual Knowledge Extraction</i>	
Ziyang Xu, Keqin Peng, Liang Ding, Dacheng Tao and Xiliang Lu	15552
<i>Take Its Essence, Discard Its Dross! Debiasing for Toxic Language Detection via Counterfactual Causal Effect</i>	
Junyu Lu, Bo Xu, Xiaokun Zhang, Kaiyuan Liu, Dongyu Zhang, Liang Yang and Hongfei Lin	15566
<i>TAPASGO: Transfer Learning towards a German-Language Tabular Question Answering Model</i>	
Dominik Andreas Kowieski, Michael Hellwig and Thomas Feilhauer	15579
<i>Target-Adaptive Consistency Enhanced Prompt-Tuning for Multi-Domain Stance Detection</i>	
Shangkang Wang and Li Pan	15585
<i>Targeted Syntactic Evaluation on the Chomsky Hierarchy</i>	
Taiga Someya, Ryo Yoshida and Yohei Oseki	15595
<i>TARIC-SLU: A Tunisian Benchmark Dataset for Spoken Language Understanding</i>	
Salima Mdhaffar, Fethi Bougares, Renato de Mori, Salah Zaiem, Mirco Ravanelli and Yannick Estève	15606
<i>TARN-VIST: Topic Aware Reinforcement Network for Visual Storytelling</i>	
Weiran Chen, Xin Li, Jiaqi Su, Guiqian Zhu, Ying Li, Yi Ji and Chunping Liu	15617
<i>Task-agnostic Distillation of Encoder-Decoder Language Models</i>	
Chen Zhang, Yang Yang, Qiuchi Li, Jingang Wang and Dawei Song	15629
<i>Task-Oriented Paraphrase Analytics</i>	
Marcel Gohsen, Matthias Hagen, Martin Potthast and Benno Stein	15640
<i>tasksource: A Large Collection of NLP tasks with a Structured Dataset Preprocessing Framework</i>	
Damien Sileo	15655
<i>Teaching Large Language Models to Translate on Low-resource Languages with Textbook Prompting</i>	
Ping Guo, Yubing Ren, Yue Hu, Yunpeng Li, Jiarui Zhang, Xingsheng Zhang and Heyan Huang	15685
<i>TECA: A Two-stage Approach with Controllable Attention Soft Prompt for Few-shot Nested Named Entity Recognition</i>	
Yuanyuan Xu, Linhai Zhang and Deyu Zhou	15698

<i>TeClass: A Human-Annotated Relevance-based Headline Classification and Generation Dataset for Telugu</i>	
Gopichand Kanumolu, Lokesh Madasu, Nirmal Surange and Manish Shrivastava ..	15711
<i>TED-EL: A Corpus for Speech Entity Linking</i>	
Silin Li, Ruoyu Song, Tianwei Lan, Zeming Liu and Yuhang Guo	15721
<i>Tell Me Again! a Large-Scale Dataset of Multiple Summaries for the Same Story</i>	
Hans Ole Hatzel and Chris Biemann	15732
<i>Temporal Knowledge Graph Reasoning with Dynamic Hypergraph Embedding</i>	
Xinyue Liu, Jianan Zhang, Chi Ma, Wenxin Liang, Bo Xu and Linlin Zong	15742
<i>Term-Driven Forward-Looking Claim Synthesis in Earnings Calls</i>	
Chung-Chi Chen and Hiroya Takamura	15752
<i>text2story: A Python Toolkit to Extract and Visualize Story Components of Narrative Text</i>	
Evelin Amorim, Ricardo Campos, Alipio Jorge, Pedro Mota and Rúben Almeida.	15761
<i>Text2Story Lusa: A Dataset for Narrative Analysis in European Portuguese News Articles</i>	
Sérgio Nunes, Alípio Mario Jorge, Evelin Amorim, Hugo Sousa, António Leal, Purificação Moura Silvano, Inês Cantante and Ricardo Campos	15773
<i>Text360Nav: 360-Degree Image Captioning Dataset for Urban Pedestrians Navigation</i>	
Chieko Nishimura, Shuhei Kurita and Yohei Seki	15783
<i>Text Filtering Classifiers for Medium-Resource Languages</i>	
Jón Daðason and Hrafn Loftsson	15789
<i>Text Style Transfer Evaluation Using Large Language Models</i>	
Phil Sidney Ostheimer, Mayank Kumar Nagda, Marius Kloft and Sophie Fellenz.	15802
<i>Text-to-Multimodal Retrieval with Bimodal Input Fusion in Shared Cross-Modal Transformer</i>	
Pranav Arora, Selen Pehlivan and Jorma Laaksonen	15823
<i>Textual Coverage of Eventive Entries in Lexical Semantic Resources</i>	
Eva Fučíková, Cristina Fernández Alcaina, Jan Hajič and Zdeňka Urešová	15835
<i>The Challenges of Creating a Parallel Multilingual Hate Speech Corpus: An Exploration</i>	
Katerina Korre, Arianna Muti and Alberto Barrón-Cedeño	15842
<i>The Contextual Variability of English Nouns: The Impact of Categorical Specificity beyond Conceptual Concreteness</i>	
Giulia Rambelli and Marianna Bolognesi	15854
<i>The Corpus AIKIA: Using Ranking Annotation for Offensive Language Detection in Modern Greek</i>	
Stella Markantonatou, Vivian Stamou, Christina Christodoulou, Georgia Apostolopoulou, Antonis Balas and George Ioannakis	15861
<i>The Distracted Ear: How Listeners Shape Conversational Dynamics</i>	
Auriane Boudin, Stéphane Rauzy, Roxane Bertrand, Magalie Ochs and Philippe Blache	15872

<i>The Effects of Pretraining in Video-Guided Machine Translation</i> Ammon Shurtz, Lawry Sorenson and Stephen D. Richardson.....	15888
<i>The ELCo Dataset: Bridging Emoji and Lexical Composition</i> Zi Yun Yang, Ziqing Zhang and Yisong Miao.....	15899
<i>The Emergence of Semantic Units in Massively Multilingual Models</i> Andrea Gregor de Varda and Marco Marelli.....	15910
<i>The Ethical Question – Use of Indigenous Corpora for Large Language Models</i> Linda Wiechetek, Flammie A. Pirinen, Børre Gaup, Trond Trosterud, Maja Lisa Kappfjell and Sjur Moshagen.....	15922
<i>The IgboAPI Dataset: Empowering Igbo Language Technologies through Multi-dialectal Enrichment</i> Chris Chinenye Emezue, Ifeoma Okoh, Chinedu Emmanuel Mbonu, Chiamaka Chukwuneke, Daisy Monika Lal, Ignatius Ezeani, Paul Rayson, Ijemma Onwuzulike, Chukwuma Onyebuchi Okeke, Gerald Okey Nweya, Bright Ikechukwu Ogbonna, Chukwuebuka Uchenna Oraegbunam, Esther Chidinma Awo-Ndubuisi and Akudo Amarachukwu Osuagwu.....	15932
<i>The Impact of Stance Object Type on the Quality of Stance Detection</i> Maxwell A. Weinzierl and Sanda M. Harabagiu.....	15942
<i>The Influence of Automatic Speech Recognition on Linguistic Features and Automatic Alzheimer’s Disease Detection from Spontaneous Speech</i> Jonathan Heitz, Gerold Schneider and Nicolas Langer.....	15955
<i>The Key Points: Using Feature Importance to Identify Shortcomings in Sign Language Recognition Models</i> Ruth M. Holmes, Ellen Rushe and Anthony Ventresque.....	15970
<i>The Low Saxon LSDC Dataset at Universal Dependencies</i> Janine Siewert and Jack Rueter.....	15976
<i>The Onomastic Repertoire of the Roman d’Alexandre (ORNARE). Designing an Integrated Digital Onomastic Tool for Medieval French Romance</i> Marta Milazzo and Giorgio Maria Di Nunzio.....	15982
<i>The Open-World Lottery Ticket Hypothesis for OOD Intent Classification</i> Yunhua Zhou, Pengyu Wang, Peiju Liu, Yuxin Wang and Xipeng Qiu.....	15988
<i>Theoretical and Empirical Advantages of Dense-Vector to One-Hot Encoding of Intent Classes in Open-World Scenarios</i> Paulo Cavalin and Claudio Santos Pinhanez.....	16000
<i>The ParCoLab Parallel Corpus and Its Extension to Four Regional Languages of France</i> Dejan Stosic, Saša Marjanović, Delphine Bernhard, Myriam Bras, Laurent Kevers, Stella Retali-Medori, Marianne Vergez-Couret and Carole Werner.....	16014
<i>The ParlaSent Multilingual Training Dataset for Sentiment Identification in Parliamentary Proceedings</i> Michal Mochtak, Peter Rupnik and Nikola Ljubešić.....	16024

<i>There's Something New about the Italian Parliament: The IPSA Corpus</i> Valentino Frasnelli and Alessio Palmero Arosio	16037
<i>The RIP Corpus of Collaborative Hypothesis-Making</i> Ella Schad, Jacky Visser and Chris Reed	16047
<i>The Role of Creaky Voice in Turn Taking and the Perception of Speaker Stance: Experiments Using Controllable TTS</i> Harm Lameris, Eva Szekely and Joakim Gustafson	16058
<i>The Role of Syntactic Span Preferences in Post-Hoc Explanation Disagreement</i> Jonathan Kamp, Lisa Beinborn and Antske Fokkens	16066
<i>The SAMER Arabic Text Simplification Corpus</i> Bashar Alhafni, Reem Hazim, Juan David Pineres Liberato, Muhamed Al Khalil and Nizar Habash	16079
<i>The Slovak Autistic and Non-Autistic Child Speech Corpus: Task-Oriented Child-Adult Interactions</i> Joanna Kruyt, Róbert Sabo, Katarína Polónyiiová, Daniela Ostatníková and Štefan Beňuš	16094
<i>The Swedish Parliament Corpus 1867 – 2022</i> Väinö Aleksy Yrjänäinen, Fredrik Mohammadi Norén, Robert Borges, Johan Jarlbrink, Lotta Åberg Brorsson, Anders P. Olsson, Pelle Snickars and Måns Magnusson	16100
<i>The Syntactic Acceptability Dataset (Preview): A Resource for Machine Learning and Linguistic Analysis of English</i> Tom S Juzek	16113
<i>The Touché23-ValueEval Dataset for Identifying Human Values behind Arguments</i> Naïlia Mirzakhmedova, Johannes Kiesel, Milad Alshomary, Maximilian Heinrich, Nicolas Handke, Xiaoni Cai, Valentin Barriere, Doratossadat Dastgheib, Omid Ghahroodi, MohammadAli SadraeiJavaheri, Ehsaneddin Asgari, Lea Kawaletz, Henning Wachsmuth and Benno Stein	16121
<i>TIGER: A Unified Generative Model Framework for Multimodal Dialogue Response Generation</i> Fanheng Kong, Peidong Wang, Shi Feng, Daling Wang and Yifei Zhang	16135
<i>TIGQA: An Expert-Annotated Question-Answering Dataset in Tigrinya</i> Hailay Kidu Teklehaymanot, Dren Fazlija, Niloy Ganguly, Gourab Kumar Patro and Wolfgang Nejdl	16142
<i>Time-aware COMET: A Commonsense Knowledge Model with Temporal Knowledge</i> Eiki Murata and Daisuke Kawahara	16162
<i>Title-based Extractive Summarization via MRC Framework</i> Hongjin Kim, Jai-Eun Kim and Harksoo Kim	16175
<i>TMFN: A Target-oriented Multi-grained Fusion Network for End-to-end Aspect-based Multimodal Sentiment Analysis</i> Di Wang, Yuzheng He, Xiao Liang, Yumin Tian, Shaofeng Li and Lin Zhao	16187

<i>To Drop or Not to Drop? Predicting Argument Ellipsis Judgments: A Case Study in Japanese</i> Yukiko Ishizuki, Tatsuki Kuribayashi, Yuichiroh Matsubayashi, Ryohei Sasano and Kentaro Inui	16198
<i>To Err Is Human, How about Medical Large Language Models? Comparing Pre-trained Language Models for Medical Assessment Errors and Reliability</i> Wen-wai Yim, Yujuan Fu, Asma Ben Abacha and Meliha Yetisgen	16211
<i>Token-length Bias in Minimal-pair Paradigm Datasets</i> Naoya Ueda, Masato Mita, Teruaki Oka and Mamoru Komachi	16224
<i>To Learn or Not to Learn: Replaced Token Detection for Learning the Meaning of Negation</i> Gunjan Bhattarai and Katrin Erk	16237
<i>ToNER: Type-oriented Named Entity Recognition with Generative Language Model</i> Guochao Jiang, Ziqin Luo, Yuchen Shi, Dixuan Wang, Jiaqing Liang and Deqing Yang	16251
<i>ToolRerank: Adaptive and Hierarchy-Aware Reranking for Tool Retrieval</i> Yuanhang Zheng, Peng Li, Wei Liu, Yang Liu, Jian Luan and Bin Wang	16263
<i>Topic Classification and Headline Generation for Maltese Using a Public News Corpus</i> Amit Kumar Chaudhary, Kurt Micallef and Claudia Borg	16274
<i>Topic-Controllable Summarization: Topic-Aware Evaluation and Transformer Methods</i> Tatiana Passali and Grigorios Tsoumakas	16282
<i>Topic Detection and Tracking with Time-Aware Document Embeddings</i> Hang Jiang, Doug Beeferman, Weiquan Mao and Deb Roy	16293
<i>TopicDiff: A Topic-enriched Diffusion Approach for Multimodal Conversational Emotion Detection</i> Jiamin Luo, Jingjing Wang and Guodong Zhou	16304
<i>Topics as Entity Clusters: Entity-based Topics from Large Language Models and Graph Neural Networks</i> Manuel V. Loureiro, Steven Derby and Tri Kurniawan Wijaya	16315
<i>To Share or Not to Share: What Risks Would Laypeople Accept to Give Sensitive Data to Differentially-Private NLP Systems?</i> Christopher Weiss, Frauke Kreuter and Ivan Habernal	16331
<i>Towards a Corpus of Spoken Maltese: Korpus tal-Malti Mitkellem, KMM</i> Alexandra (Sandra) Vella, Sarah Agius, Aiden Williams and Claudia Borg	16343
<i>Towards a Danish Semantic Reasoning Benchmark - Compiled from Lexical-Semantic Resources for Assessing Selected Language Understanding Capabilities of Large Language Models</i> Bolette Pedersen, Nathalie Sørensen, Sussi Olsen, Sanni Nimb and Simon Gray ..	16353
<i>Towards a Framework for Evaluating Explanations in Automated Fact Verification</i> Neema Kotonya and Francesca Toni	16364

<i>Towards Algorithmic Fidelity: Mental Health Representation across Demographics in Synthetic vs. Human-generated Data</i>	
Shinka Mori, Oana Ignat, Andrew Lee and Rada Mihalcea	16378
<i>Towards an Ideal Tool for Learner Error Annotation</i>	
Špela Arhar Holdt, Tomaž Erjavec, Iztok Kosem and Elena Volodina	16392
<i>Towards Answering Health-related Questions from Medical Videos: Datasets and Approaches</i>	
Deepak Gupta, Kush Attal and Dina Demner-Fushman	16399
<i>Towards a Unified Taxonomy of Deep Syntactic Relations</i>	
Kira Droganova and Daniel Zeman	16412
<i>Towards Autonomous Tool Utilization in Language Models: A Unified, Efficient and Scalable Framework</i>	
Zhi Li, Yicheng Li, Hequan Ye and Yin Zhang	16422
<i>Towards a Zero-Data, Controllable, Adaptive Dialog System</i>	
Dirk Văth, Lindsey Vanderlyn and Ngoc Thang Vu	16433
<i>Towards Building the LEMI Readability Platform for Children's Literature in the Romanian Language</i>	
Madalina Chitez, Mihai Dascalu, Aura Cristina Udrea, Cosmin Strilețchi, Karla Csürös, Roxana Rogobete and Alexandru Oravițan	16450
<i>Towards Comprehensive Language Analysis for Clinically Enriched Spontaneous Dialogue</i>	
Baris Karacan, Ankit Aich, Avery Quynh, Amy Pinkham, Philip Harvey, Colin Depp and Natalie Parde	16457
<i>Towards Cost-effective Multi-style Conversations: A Pilot Study in Task-oriented Dialogue Generation</i>	
Tiziano Labruna and Bernardo Magnini	16473
<i>Towards Dog Bark Decoding: Leveraging Human Speech Processing for Automated Bark Classification</i>	
Artem Abzaliev, Humberto Perez-Espinosa and Rada Mihalcea	16480
<i>Towards Equitable Natural Language Understanding Systems for Dialectal Cohorts: Debiasing Training Data</i>	
Khadije Abboud and Gokmen Oz	16487
<i>Towards Explainability and Fairness in Swiss Judgement Prediction: Benchmarking on a Multilingual Dataset</i>	
Santosh T.Y.S.S., Nina Baumgartner, Matthias Stürmer, Matthias Grabmair and Joel Niklaus	16500
<i>Towards Few-shot Entity Recognition in Document Images: A Graph Neural Network Approach Robust to Image Manipulation</i>	
Prashant Krishnan, Zilong Wang, Yangkun Wang and Jingbo Shang	16514
<i>Towards Generalizable and Faithful Logic Reasoning over Natural Language via Resolution Refutation</i>	
Zhouhao Sun, Xiao Ding, Li Du, Bibo Cai, Jinglong Gao, Ting Liu and Bing Qin	16527

<i>Towards Graph-hop Retrieval and Reasoning in Complex Question Answering over Textual Database</i>	
Minjun Zhu, Yixuan Weng, Shizhu He, Kang Liu, Haifeng Liu, Yang jun Jun and Jun Zhao	16539
<i>Towards Human-aligned Evaluation for Linear Programming Word Problems</i>	
Linzi Xing, Xinglu Wang, Yuxi Feng, Zhenan Fan, Jing Xiong, Zhijiang Guo, Xiaojin Fu, Rindra Ramamonjison, Mahdi Mostajabdaveh, Xiongwei Han, Zirui Zhou and Yong Zhang	16550
<i>Towards Human-Like Machine Comprehension: Few-Shot Relational Learning in Visually-Rich Documents</i>	
Hao Wang, Tang Li, Chenhui Chu, Rui Wang and Pinpin Zhu	16557
<i>Towards More Realistic Chinese Spell Checking with New Benchmark and Specialized Expert Model</i>	
Yue Wang, Zilong Zheng, Juntao Li, Zhihui Liu, Jinxiong Chang, Qishen Zhang, Zhongyi Liu, Guannan Zhang and Min Zhang	16570
<i>Towards Multi-modal Sarcasm Detection via Disentangled Multi-grained Multi-modal Distilling</i>	
Zhihong Zhu, Xuxin Cheng, Guimin Hu, Yaowei Li, Zhiqi Huang and Yuexian Zou	16581
<i>Towards Realistic Few-Shot Relation Extraction: A New Meta Dataset and Evaluation</i>	
Fahmida Alam, Md Asiful Islam, Robert Vacareanu and Mihai Surdeanu	16592
<i>Towards Robust Evidence-Aware Fake News Detection via Improving Semantic Perception</i>	
Yike Wu, Yang Xiao, Mengting Hu, Mengying Liu, Pengcheng Wang and Mingming Liu	16607
<i>Towards Robust In-Context Learning for Machine Translation with Large Language Models</i>	
Shaolin Zhu, Menglong Cui and Deyi Xiong	16619
<i>Towards Robust Temporal Activity Localization Learning with Noisy Labels</i>	
Daizong Liu, Xiaoye Qu, Xiang Fang, Jianfeng Dong, Pan Zhou, Guoshun Nan, Keke Tang, Wanlong Fang and Yu Cheng	16630
<i>Towards Semantic Tagging for Irish</i>	
Tim Czerniak and Elaine Uí Dhonnchadha	16643
<i>Towards Standardized Annotation and Parsing for Korean FrameNet</i>	
Yige Chen, Jae Ihn, KyungTae Lim and Jungyeul Park	16653
<i>Towards the WhAP Corpus: A Resource for the Study of Italian on WhatsApp</i>	
Ilaria Fiorentini, Marco Forlano and Nicholas Nese	16659
<i>Towards Understanding the Relationship between In-context Learning and Compositional Generalization</i>	
Sungjun Han and Sebastian Padó	16664
<i>Towards Universal Dependencies for Ancash Quechua</i>	
Johanna Cordova	16680
<i>TP-Link: Fine-grained Pre-Training for Text-to-SQL Parsing with Linking Information</i>	
Ziqiang Liu, Shujie Li, Zefeng Cai, Xiangyu Li, Yunshui Li, Chengming Li, Xiping Hu, Ruifeng Xu and Min Yang	16686

<i>Training BERT Models to Carry over a Coding System Developed on One Corpus to Another</i> Dalma Galambos and Pal Zsamboki	16698
<i>TransCoder: Towards Unified Transferable Code Representation Learning Inspired by Human Skills</i> Qiushi Sun, Nuo Chen, Jianing Wang, Ming Gao and Xiang Li	16713
<i>TransERR: Translation-based Knowledge Graph Embedding via Efficient Relation Rotation</i> Jiang Li, Xiangdong Su, Fujun Zhang and Guanglai Gao	16727
<i>Transfer Fine-tuning for Quality Estimation of Text Simplification</i> Yuki Hironaka, Tomoyuki Kajiwara and Takashi Ninomiya	16738
<i>Transferring BERT Capabilities from High-Resource to Low-Resource Languages Using Vocabulary Matching</i> Piotr Rybak	16745
<i>Transformer-based Joint Modelling for Automatic Essay Scoring and Off-Topic Detection</i> Sourya Dipta Das, Yash A. Vadi and Kuldeep Yadav	16751
<i>Transformer-based Swedish Semantic Role Labeling through Transfer Learning</i> Dana Dannélls, Richard Johansson and Lucy Yang Buhr	16762
<i>Transformers for Bridging Persian Dialects: Transliteration Model for Tajiki and Iranian Scripts</i> MohammadAli SadraeiJavaheri, Ehsaneddin Asgari and Hamid Reza Rabiee	16770
<i>Tree-Instruct: A Preliminary Study of the Intrinsic Relationship between Complexity and Alignment</i> Yingxiu Zhao, Bowen Yu, Binyuan Hui, Haiyang Yu, Minghao Li, Fei Huang, Nevin L. Zhang and Yongbin Li	16776
<i>TRELM: Towards Robust and Efficient Pre-training for Knowledge-Enhanced Language Models</i> Junbing Yan, Chengyu Wang, Taolin Zhang, Xiaofeng He, Jun Huang, Wei Zhang, Longtao Huang and Hui Xue	16790
<i>Tricking LLMs into Disobedience: Formalizing, Analyzing, and Detecting Jailbreaks</i> Abhinav Sukumar Rao, Atharva Roshan Naik, Sachin Vashistha, Somak Aditya and Monojit Choudhury	16802
<i>Triple-R: Automatic Reasoning for Fact Verification Using Language Models</i> Mohammadamin Kanaani	16831
<i>Triples-to-isiXhosa (T2X): Addressing the Challenges of Low-Resource Agglutinative Data-to-Text Generation</i> Francois Meyer and Jan Buys	16841
<i>Trustworthiness and Self-awareness in Large Language Models: An Exploration through the Think-Solve-Verify Framework</i> Zhendong Liu, Changhong Xia, Wei He and Chongjun Wang	16855
<i>Tug-of-War between Knowledge: Exploring and Resolving Knowledge Conflicts in Retrieval-Augmented Language Models</i> Zhuoran Jin, Pengfei Cao, Yubo Chen, Kang Liu, Xiaojian Jiang, Jiexin Xu, Li Qiuxia and Jun Zhao	16867

<i>TunArTTS: Tunisian Arabic Text-To-Speech Corpus</i>	
Imen Laouirine, Rami Kammoun and Fethi Bougares	16879
<i>TweetTER: A Benchmark for Target Entity Retrieval on Twitter without Knowledge Bases</i>	
Kiamehr Rezaee, Jose Camacho-Collados and Mohammad Taher Pilehvar	16890
<i>Two Counterexamples to Tokenization and the Noiseless Channel</i>	
Marco Cognaetta, Vilém Zouhar, Sangwhan Moon and Naoaki Okazaki	16897
<i>Typos Correction Training against Misspellings from Text-to-Text Transformers</i>	
Guicai Xie, Ke Zhang, Lei Duan, Wei Zhang and Zeqian Huang	16907
<i>UCxn: Typologically-Informed Annotation of Constructions Atop Universal Dependencies</i>	
Leonie Weissweiler, Nina Böbel, Kirian Guiller, Santiago Herrera, Wesley Samuel Scivetti, Arthur Lorenzi, Nurit Melnik, Archana Bhatia, Hinrich Schütze, Lori Levin, Amir Zeldes, Joakim Nivre, William Croft and Nathan Schneider	16919
<i>UDMorph: Morphosyntactically Tagged UD Corpora</i>	
Maarten Janssen	16933
<i>UkraiNER: A New Corpus and Annotation Scheme towards Comprehensive Entity Recognition</i>	
Lauriane Aufrant and Lucie Chasseur	16941
<i>UMTIT: Unifying Recognition, Translation, and Generation for Multimodal Text Image Translation</i>	
Liqiang Niu, Fandong Meng and Jie Zhou	16953
<i>Uncertainty-Aware Cross-Modal Alignment for Hate Speech Detection</i>	
Chuanpeng Yang, Fuqing Zhu, Yaxin Liu, Jizhong Han and Songlin Hu	16973
<i>Uncovering Agendas: A Novel French & English Dataset for Agenda Detection on Social Media</i>	
Gregorios Katsios, Ning Sa, Ankita Bhaumik and Tomek Strzalkowski	16984
<i>Uncovering the Potential of ChatGPT for Discourse Analysis in Dialogue: An Empirical Study</i>	
Yaxin Fan, Feng Jiang, Peifeng Li and Haizhou Li	16998
<i>Understanding How Positional Encodings Work in Transformer Model</i>	
Taro Miyazaki, Hideya Mino and Hiroyuki Kaneko	17011
<i>Unicode Normalization and Grapheme Parsing of Indic Languages</i>	
Nazmuddoha Ansary, Quazi Adibur Rahman Adib, Tahsin Reasat, Asif Shahriyar Sushmit, Ahmed Imtiaz Humayun, Sazia Mehnaz, Kanij Fatema, Mohammad Mamun Or Rashid and Farig Sadeque	17019
<i>Unifying Latent and Lexicon Representations for Effective Video-Text Retrieval</i>	
Haowei Liu, Yaya Shi, Haiyang Xu, Chunfeng Yuan, Qinghao Ye, Chenliang Li, Ming Yan, Ji Zhang, Fei Huang, Bing Li and Weiming Hu	17031
<i>UniPCM: Universal Pre-trained Conversation Model with Task-aware Automatic Prompt</i>	
Yucheng Cai, Wentao Ma, Yuchuan Wu, Shuzheng Si, Yuan Shao, Zhijian Ou and Yongbin Li	17042

<i>UniPSDA: Unsupervised Pseudo Semantic Data Augmentation for Zero-Shot Cross-Lingual Natural Language Understanding</i>	
Dongyang Li, Taolin Zhang, Jiali Deng, Longtao Huang, Chengyu Wang, Xiaofeng He and Hui Xue	17062
<i>UniRetriever: Multi-task Candidates Selection for Various Context-Adaptive Conversational Retrieval</i>	
Hongru Wang, Boyang Xue, Baohang Zhou, Rui Wang, Fei Mi, Weichao Wang, Yasheng Wang and Kam-Fai Wong	17074
<i>Universal Anaphora: The First Three Years</i>	
Massimo Poesio, Maciej Ogródniczuk, Vincent Ng, Sameer Pradhan, Juntao Yu, Nafise Sادات Moosavi, Silviu Paun, Amir Zeldes, Anna Nedoluzhko, Michal Novák, Martin Popel, Zdeněk Žabokrtský and Daniel Zeman	17087
<i>Universal Dependencies: Extensions for Modern and Historical German</i>	
Stefanie Dipper, Cora Haiber, Anna Maria Schröter, Alexandra Wiemann and Maike Brinkschulte	17101
<i>Universal Dependencies for Learner Russian</i>	
Alla Rozovskaya	17112
<i>Unleashing the Power of Imbalanced Modality Information for Multi-modal Knowledge Graph Completion</i>	
Yichi Zhang, Zhuo Chen, Lei Liang, Huajun Chen and Wen Zhang	17120
<i>Unlocking Instructive In-Context Learning with Tabular Prompting for Relational Triple Extraction</i>	
Guozheng Li, Wenjun Ke, Peng Wang, Zijie Xu, Ke Ji, Jiajun Liu, Ziyu Shang and Qiqing Luo	17131
<i>Unmasking Biases: Exploring Gender Bias in English-Catalan Machine Translation through Tokenization Analysis and Novel Dataset</i>	
Audrey Mash, Carlos Escolano, Aleix Sant, Maite Melero and Francesca de Luca Fornaciari	17144
<i>Unpacking Bias: An Empirical Study of Bias Measurement Metrics, Mitigation Algorithms, and Their Interactions</i>	
Felipe Bravo-Marquez and Maria Jose Zambrano	17154
<i>Unraveling Spontaneous Speech Dimensions for Cross-Corpus ASR System Evaluation for French</i>	
Solene Virginie Evain, Solange Rossato and François Portet	17165
<i>Unsupervised Grouping of Public Procurement Similar Items: Which Text Representation Should I Use?</i>	
Pedro P. V. Brum, Mariana O. Silva, Gabriel P. Oliveira, Lucas G. L. Costa, Anisio Lacerda and Gisele Pappa	17176
<i>Untangle the KNOT: Interweaving Conflicting Knowledge and Reasoning Skills in Large Language Models</i>	
Yantao Liu, Zijun Yao, Xin Lv, Yuchen Fan, Shulin Cao, Jifan Yu, Lei Hou and Juanzi Li	17186

<i>Unveiling Project-Specific Bias in Neural Code Models</i> Zhiming Li, Yanzhou Li, Tianlin Li, Mengnan Du, Bozhi Wu, Yushi Cao, Junzhe Jiang and Yang Liu	17205
<i>Unveiling Strengths and Weaknesses of NLP Systems Based on a Rich Evaluation Corpus: The Case of NER in French</i> Alice Millour, Yoann Dupont, Karen Fort and Liam Duignan	17217
<i>Unveiling Vulnerability of Self-Attention</i> Khai Jiet Liong, Hongqiu Wu and Hai Zhao	17225
<i>UQA: Corpus for Urdu Question Answering</i> Samee Arif, Sualeha Farid, Awais Athar and Agha Ali Raza	17237
<i>UrduMASD: A Multimodal Abstractive Summarization Dataset for Urdu</i> Ali Faheem, Faizad Ullah, Muhammad Sohaib Ayub and Asim Karim	17245
<i>User Guide for KOTE: Korean Online That-gul Emotions Dataset</i> Duyoung Jeon, Junho Lee and Cheongtag Kim	17254
<i>Using Bibliodata LODification to Create Metadata-Enriched Literary Corpora in Line with FAIR Principles</i> Agnieszka Karlinska, Cezary Rosiński, Marek Kubis, Patryk Hubar and Jan Wiczorek	17271
<i>Using Persuasive Writing Strategies to Explain and Detect Health Misinformation</i> Danial Kamali, Joseph D. Romain, Huiyi Liu, Wei Peng, Jingbo Meng and Parisa Kordjamshidi	17285
<i>Using Pre-Trained Language Models in an End-to-End Pipeline for Antithesis Detection</i> Ramona Kühn, Khoulood Saadi, Jelena Mitrović and Michael Granitzer	17310
<i>Using Speech Technology to Test Theories of Phonetic and Phonological Typology</i> Anisia Popescu, Lori Lamel and Ioana Vasilescu	17321
<i>Utilizing Local Hierarchy with Adversarial Training for Hierarchical Text Classification</i> Zihan Wang, Peiyi Wang and Houfeng Wang	17326
<i>Utilizing Longer Context than Speech Bubbles in Automated Manga Translation</i> Hiroto Kaino, Soichiro Sugihara, Tomoyuki Kajiwara, Takashi Ninomiya, Joshua B. Tanner and Shonosuke Ishiwatari	17337
<i>UzbekVerbDetection: Rule-based Detection of Verbs in Uzbek Texts</i> Maksud Sharipov, Elmurod Kuriyozov, Ollabergan Yuldashev and Ogabek Sobirov .	17343
<i>Validating and Exploring Large Geographic Corpora</i> Jonathan Dunn	17348
<i>Verbing Weirds Language (Models): Evaluation of English Zero-Derivation in Five LLMs</i> David R. Mortensen, Valentina Izrailevitch, Yunze Xiao, Hinrich Schütze and Leonie Weisweiler	17359

<i>VietMed: A Dataset and Benchmark for Automatic Speech Recognition of Vietnamese in the Medical Domain</i> Khai Le-Duc	17365
<i>VI-OOD: A Unified Framework of Representation Learning for Textual Out-of-distribution Detection</i> Li-Ming Zhan, Bo Liu and Xiao-Ming Wu	17371
<i>Visual-Linguistic Dependency Encoding for Image-Text Retrieval</i> Wenxin Guo, Lei Zhang, Kun Zhang, Yi Liu and Zhendong Mao	17384
<i>Visual-Textual Entailment with Quantities Using Model Checking and Knowledge Injection</i> Nobuyuki Iokawa and Hitomi Yanaka	17397
<i>Vygotsky Distance: Measure for Benchmark Task Similarity</i> Maxim K. Surkov and Ivan P. Yamshchikov	17409
<i>WaCadie: Towards an Acadian French Corpus</i> Jeremy Robichaud and Paul Cook	17421
<i>Well Begun Is Half Done: An Implicitly Augmented Generative Framework with Distribution Modification for Hierarchical Text Classification</i> Huawen Feng, Jingsong Yan, Junlong Liu, Junhao Zheng and Qianli Ma	17433
<i>What Are the Implications of Your Question? Non-Information Seeking Question-Type Identification in CNN Transcripts</i> Yao Sun, Anastasiia Tatlubaeva, Zhihan Li and Chester Palen-Michel	17444
<i>What Can Diachronic Contexts and Topics Tell Us about the Present-Day Compositionality of English Noun Compounds?</i> Samin Mahdizadeh Sani, Malak Rassem, Chris W. Jenkins, Filip Miletic and Sabine Schulte im Walde	17449
<i>What Do Transformers Know about Government?</i> Jue Hou, Anisia Katinskaia, Lari Kotilainen, Sathianpong Trangcasanchai, Anh-Duc Vu and Roman Yangarber	17459
<i>What Factors Influence LLMs' Judgments? A Case Study on Question Answering</i> Lei Chen, Bobo Li, Li Zheng, Haining Wang, Zixiang Meng, Runfeng Shi, Hao Fei, Jun Zhou, Fei Li, Chong Teng and Donghong Ji	17473
<i>What Happens to a Dataset Transformed by a Projection-based Concept Removal Method?</i> Richard Johansson	17486
<i>What Has LeBenchmark Learnt about French Syntax?</i> Zdravko Dugonjic, Adrien Pupier, Benjamin Lecouteux and Maximin Coavoux	17493
<i>What Is Needed for Intra-document Disambiguation of Math Identifiers?</i> Takuto Asakura and Yusuke Miyao	17500
<i>When Argumentation Meets Cohesion: Enhancing Automatic Feedback in Student Writing</i> Yuning Ding, Omid Kashefi, Swapna Somasundaran and Andrea Horbach	17513

<i>When Cohesion Lies in the Embedding Space: Embedding-Based Reference-Free Metrics for Topic Segmentation</i>	Iacopo Ghinassi, Lin Wang, Chris Newell and Matthew Purver	17525
<i>When Do "More Contexts" Help with Sarcasm Recognition?</i>	Ojas Nimase and Sanghyun Hong	17537
<i>When Your Cousin Has the Right Connections: Unsupervised Bilingual Lexicon Induction for Related Data-Imbalanced Languages</i>	Niyati Bafna, Cristina España-Bonet, Josef van Genabith, Benoît Sagot and Rachel Bawden	17544
<i>Which Sense Dominates Multisensory Semantic Understanding? A Brain Decoding Study</i>	Dandan Huang, Lu Cao, Zhenting Li and Yue Zhang	17557
<i>Who Did You Blame When Your Project Failed? Designing a Corpus for Presupposition Generation in Cross-Examination Dialogues</i>	Maria Francis, Julius Steuer, Dietrich Klakow and Volha Petukhova	17564
<i>Who Is Bragging More Online? A Large Scale Analysis of Bragging in Social Media</i>	Mali Jin, Daniel Preotiuc-Pietro, A. Seza Dođruöz and Nikolaos Aletras	17575
<i>Who Said What: Formalization and Benchmarks for the Task of Quote Attribution</i>	Wenjie Zhong, Jason Naradowsky, Hiroya Takamura, Ichiro Kobayashi and Yusuke Miyao	17588
<i>Why Voice Biomarkers of Psychiatric Disorders Are Not Used in Clinical Practice? Deconstructing the Myth of the Need for Objective Diagnosis</i>	Vincent P. Martin and Jean-Luc Rouas	17603
<i>WikiFactDiff: A Large, Realistic, and Temporally Adaptable Dataset for Atomic Factual Knowledge Update in Causal Language Models</i>	Hichem Ammar Khodja, Frederic Bechet, Quentin Brabant, Alexis Nasr and Gwénolé Lecorvé	17614
<i>WikiSplit++: Easy Data Refinement for Split and Rephrase</i>	Hayato Tsukagoshi, Tsutomu Hirao, Makoto Morishita, Katsuki Chousa, Ryohei Sasano and Koichi Takeda	17625
<i>Willkommens-Merkel, Chaos-Johnson, and Tore-Klose: Modeling the Evaluative Meaning of German Personal Name Compounds</i>	Annerose Eichel, Tana Deeg, Andre Blessing, Milena Belosevic, Sabine Arndt-Lappe and Sabine Schulte im Walde	17637
<i>WkNER: Enhancing Named Entity Recognition with Word Segmentation Constraints and kNN Retrieval</i>	Yanchun Li, Senlin Deng, Dongsu Shen, Shujuan Tian and Saiqin Long	17651
<i>Word-Aware Modality Stimulation for Multimodal Fusion</i>	Shuheite Tateishi, Makoto Nakatsuji and Yasuhito Osugi	17664
<i>Word-level Commonsense Knowledge Selection for Event Detection</i>	Shuai Yang, Yu Hong, Shiming He, Qingting Xu and Jianmin Yao	17675

<i>WordNet under Scrutiny: Dictionary Examples in the Era of Large Language Models</i> Fatemah Yousef Almeman, Steven Schockaert and Luis Espinosa Anke	17683
<i>WorldValuesBench: A Large-Scale Benchmark Dataset for Multi-Cultural Value Awareness of Language Models</i> Wenlong Zhao, Debanjan Mondal, Niket Tandon, Danica Dillion, Kurt Gray and Yuling Gu	17696
<i>Would You Like to Make a Donation? A Dialogue System to Persuade You to Donate</i> Yuhan Song and Houfeng Wang	17707
<i>WW-CSL: A New Dataset for Word-Based Wearable Chinese Sign Language Detection</i> Fan Xu, Kai Liu, Yifeng Yang and Keyu Yan.....	17718
<i>XAI-Attack: Utilizing Explainable AI to Find Incorrectly Learned Patterns for Black-Box Adversarial Example Creation</i> Markus Bayer, Markus Neiczer, Maximilian Samsinger, Björn Buchhold and Christian Reuter	17725
<i>XATU: A Fine-grained Instruction-based Benchmark for Explainable Text Updates</i> Haopeng Zhang, Hayate Iso, Sairam Gurajada and Nikita Bhutani	17739
<i>XVD: Cross-Vocabulary Differentiable Training for Generative Adversarial Attacks</i> Tom Roth, Inigo Jauregi Unanue, Alsharif Abuadbbba and Massimo Piccardi.....	17753
<i>Your Stereotypical Mileage May Vary: Practical Challenges of Evaluating Biases in Multiple Languages and Cultural Contexts</i> Karen Fort, Laura Alonso Alemany, Luciana Benotti, Julien Bezançon, Claudia Borg, Marthese Borg, Yongjian Chen, Fanny DuceL, Yoann Dupont, Guido Ivetta, Zhijian Li, Margot Mieskes, Marco Naguib, Yuyan Qian, Matteo Radaelli, Wolfgang S. Schmeisser-Nieto, Emma Raimundo Schulz, Thiziri Saci, Sarah Saidi, Javier Torroba Marchante, Shilin Xie, Sergio E. Zanotto and Aurélie Névéal	17764
<i>ZAEBUC-Spoken: A Multilingual Multidialectal Arabic-English Speech Corpus</i> Injy Hamed, Fadhl Eryani, David Palfreyman and Nizar Habash.....	17770
<i>ZeLa: Advancing Zero-Shot Multilingual Semantic Parsing with Large Language Models and Chain-of-Thought Strategies</i> Truong Dinh Do, Phuong Minh Nguyen and Minh Nguyen.....	17783
<i>ZenPropaganda: A Comprehensive Study on Identifying Propaganda Techniques in Russian Coronavirus-Related Media</i> Anton Chernyavskiy, Svetlana Shomova, Irina Dushakova, Ilya Kiriya and Dmitry Ilvovsky	17795
<i>Zero- and Few-Shot Prompting with LLMs: A Comparative Study with Fine-tuned Models for Bangla Sentiment Analysis</i> Md. Arid Hasan, Shudipta Das, Afiyat Anjum, Firoj Alam, Anika Anjum, Avijit Sarker and Sheak Rashed Haider Noori	17808
<i>Zero-shot Cross-lingual Automated Essay Scoring</i> Junyi He and Xia Li	17819

Zero-Shot Cross-Lingual Document-Level Event Causality Identification with Heterogeneous Graph Contrastive Transfer Learning

Zhitao He, Pengfei Cao, Zhuoran Jin, Yubo Chen, Kang Liu, Zhiqiang Zhang, Mengshu Sun and Jun Zhao 17833

Zero-shot Event Detection Using a Textual Entailment Model as an Enhanced Annotator

Ziqian Zeng, Runyu Wu, Yuxiang Xiao, Xiaoda Zhong, Hanlin Wang, Zhengdong Lu and Huiping Zhuang 17851

Zero-shot Learning for Multilingual Discourse Relation Classification

Eleni Metheniti, Philippe Muller, Chloé Braud and Margarita Hernández Casas 17858

Zero-Shot Spoken Language Understanding via Large Language Models: A Preliminary Study

Zhihong Zhu, Xuxin Cheng, Hao An, Zhichang Wang, Dongsheng Chen and Zhiqi Huang
17877