

# **18th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2024)**

Findings of EACL 2024

St. Julian's, Malta  
21 - 22 March 2024

Part 1 of 3

ISBN: 978-1-7138-9447-6

**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 the Association for Computational Linguistics  
All rights reserved.

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

For permission requests, please contact the Association for Computational Linguistics  
at the address below.

Association for Computational Linguistics  
209 N. Eighth Street  
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006

Fax: 1-570-476-0860

[acl@aclweb.org](mailto:acl@aclweb.org)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2633  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

## Table of Contents

<i>Chem-FINESE: Validating Fine-Grained Few-shot Entity Extraction through Text Reconstruction</i> Qingyun Wang, Zixuan Zhang, Hongxiang Li, Xuan Liu, Jiawei Han, Huimin Zhao and Heng Ji	
<i>GPTs Are Multilingual Annotators for Sequence Generation Tasks</i> Juhwan Choi, Eunju Lee, Kyohoon Jin and YoungBin Kim	17
<i>Next Visit Diagnosis Prediction via Medical Code-Centric Multimodal Contrastive EHR Modelling with Hierarchical Regularisation</i> Heejoon Koo	41
<i>FlexiQA: Leveraging LLM’s Evaluation Capabilities for Flexible Knowledge Selection in Open-domain Question Answering</i> Yuhan Chen, Shuqi Li and Rui Yan	56
<i>Hyper-BTS Dataset: Scalability and Enhanced Analysis of Back Transcription (BTS) for ASR Post-Processing</i> Chanjun Park, Jaehyung Seo, Seolhwa Lee, Junyoung Son, Hyeonseok Moon, Sugyeong Eo, Chanhee Lee and Heuiseok Lim	67
<i>ParrotTTS: Text-to-speech synthesis exploiting disentangled self-supervised representations</i> Neil Shah, Saiteja Kosgi, Vishal Tambrhalli, Neha S, Anil Kumar Nelakanti and Vineet Gandhi	79
<i>NavHint: Vision and Language Navigation Agent with a Hint Generator</i> Yue Zhang, Quan Guo and Parisa Kordjamshidi	92
<i>Text or Image? What is More Important in Cross-Domain Generalization Capabilities of Hate Meme Detection Models?</i> Piush Aggarwal, Jawar Mehrabian, Weigang Huang, Özge Alacam and Torsten Zesch	104
<i>Where are we Still Split on Tokenization?</i> Rob Van Der Goot	118
<i>A Methodology for Generative Spelling Correction via Natural Spelling Errors Emulation across Multiple Domains and Languages</i> Nikita Martynov, Mark Baushenko, Anastasia Kozlova, Katerina Kolomeytseva, Aleksandr Abramov and Alena Fenogenova	138
<i>How Does In-Context Learning Help Prompt Tuning?</i> Simeng Sun, Yang Liu, Dan Iter, Chenguang Zhu and Mohit Iyyer	156
<i>Large Language Models for Psycholinguistic Plausibility Pretesting</i> Samuel Joseph Amouyal, Aya Meltzer-Asscher and Jonathan Berant	166
<i>Modeling Aspect Sentiment Coherency via Local Sentiment Aggregation</i> Heng Yang and Ke Li	182
<i>An Examination of the Robustness of Reference-Free Image Captioning Evaluation Metrics</i> Saba Ahmadi and Aishwarya Agrawal	196
<i>Barriers to Effective Evaluation of Simultaneous Interpretation</i> Shira Wein, Te I, Colin Cherry, Juraj Juraska, Dirk Padfield and Wolfgang Macherey	209

<i>Inconsistent dialogue responses and how to recover from them</i>	
Mian Zhang, Lifeng Jin, Linfeng Song, Haitao Mi and Dong Yu .....	220
<i>MUG: Interactive Multimodal Grounding on User Interfaces</i>	
Tao Li, Gang Li, Jingjie Zheng, Purple Wang and Yang Li .....	231
<i>PRILoRA: Pruned and Rank-Increasing Low-Rank Adaptation</i>	
Nadav Benedek and Lior Wolf .....	252
<i>Revamping Multilingual Agreement Bidirectionally via Switched Back-translation for Multilingual Neural Machine Translation</i>	
Hongyuan Lu, Haoyang Huang, Dongdong Zhang, Furu Wei and Wai Lam .....	264
<i>mPLM-Sim: Better Cross-Lingual Similarity and Transfer in Multilingual Pretrained Language Models</i>	
Peiqin Lin, Chengzhi Hu, Zheyu Zhang, Andre Martins and Hinrich Schuetze .....	276
<i>OYXOY: A Modern NLP Test Suite for Modern Greek</i>	
Konstantinos Kogkalidis, Stergios Chatzikyriakidis, Eirini Chrysovalantou Giannikouri, Vasiliki Katsouli, Christina Klironomou, Christina Koula, Dimitris Papadakis, Thelka Pasparaki, Erofilia Psaltaki, Efthymia Sakellariou and Charikleia Soupiona .....	311
<i>A Comprehensive Evaluation of Inductive Reasoning Capabilities and Problem Solving in Large Language Models</i>	
Chen Bowen, Rune Sætre and Yusuke Miyao .....	323
<i>Towards efficient self-supervised representation learning in speech processing</i>	
Luis Lugo and Valentin Vielzeuf .....	340
<i>Improving Cross-Domain Low-Resource Text Generation through LLM Post-Editing: A Programmer-Interpreter Approach</i>	
Zhuang Li, Levon Haroutunian, Raj Tumuluri, Philip R. Cohen and Reza Haf .....	347
<i>Noise Contrastive Estimation-based Matching Framework for Low-Resource Security Attack Pattern Recognition</i>	
Tu Nguyen, Nedim Šrndić and Alexander Neth .....	355
<i>Large Language Models for Scientific Information Extraction: An Empirical Study for Virology</i>	
Mahsa Shamsabadi, Jennifer D’Souza and Sören Auer .....	374
<i>Re3val: Reinforced and Reranked Generative Retrieval</i>	
EuiYul Song, Sangryul Kim, Haeju Lee, Joonkee Kim and James Thorne .....	393
<i>Entity Linking in the Job Market Domain</i>	
Mike Zhang, Rob Van Der Goot and Barbara Plank .....	410
<i>(Chat)GPT v BERT Dawn of Justice for Semantic Change Detection</i>	
Francesco Periti, Haim Dubossarsky and Nina Tahmasebi .....	420
<i>Towards Unified Uni- and Multi-modal News Headline Generation</i>	
Mateusz Krubiński and Pavel Pecina .....	437
<i>On the Relationship between Sentence Analogy Identification and Sentence Structure Encoding in Large Language Models</i>	
Thilini Wijesiriwardene, Ruwan Wickramarachchi, Aishwarya Naresh Reganti, Vinija Jain, Aman Chadha, Amit P. Sheth and Amitava Das .....	451
<i>Contextualization Distillation from Large Language Model for Knowledge Graph Completion</i>	
Dawei Li, Zhen Tan, Tianlong Chen and Huan Liu .....	458

<i>Differentially Private Natural Language Models: Recent Advances and Future Directions</i> Lijie Hu, Ivan Habernal, Lei Shen and Di Wang .....	478
<i>Learning to Compare Financial Reports for Financial Forecasting</i> Ross Koval, Nicholas Andrews and Xifeng Yan .....	500
<i>Arukikata Travelogue Dataset with Geographic Entity Mention, Coreference, and Link Annotation</i> Shohei Higashiyama, Hiroki Ouchi, Hiroki Teranishi, Hiroyuki Otomo, Yusuke Ide, Aitaro Yamamoto, Hiroyuki Shindo, Yuki Matsuda, Shoko Wakamiya, Naoya Inoue, Ikuya Yamada and Taro Watanabe.....	513
<i>Knowledge Generation for Zero-shot Knowledge-based VQA</i> Rui Cao and Jing Jiang .....	533
<i>Simple Temperature Cool-down in Contrastive Framework for Unsupervised Sentence Representation Learning</i> Yoo Hyun Jeong, Myeong Soo Han and Dong-Kyu Chae .....	550
<i>Bootstrap Your Own PLM: Boosting Semantic Features of PLMs for Unsupervised Contrastive Learning</i> Yoo Hyun Jeong, Myeong Soo Han and Dong-Kyu Chae .....	560
<i>Personalized Abstractive Summarization by Tri-agent Generation Pipeline</i> Wen Xiao, Yujia Xie, Giuseppe Carenini and Pengcheng He .....	570
<i>Revisiting the Markov Property for Machine Translation</i> Cunxiao Du, Hao Zhou, Zhaopeng Tu and Jing Jiang .....	582
<i>Reward Engineering for Generating Semi-structured Explanation</i> Jiuzhou Han, Wray Buntine and Ehsan Shareghi.....	589
<i>Towards Context-Based Violence Detection: A Korean Crime Dialogue Dataset</i> Minju Kim, Heuiyeen Yeen and Myoung-Wan Koo .....	603
<i>Capturing the Relationship Between Sentence Triplets for LLM and Human-Generated Texts to Enhance Sentence Embeddings</i> Na Min An, Sania Waheed and James Thorne .....	624
<i>Harmonizing Code-mixed Conversations: Personality-assisted Code-mixed Response Generation in Dialogues</i> Shivani Kumar and Tanmoy Chakraborty .....	639
<i>Morality is Non-Binary: Building a Pluralist Moral Sentence Embedding Space using Contrastive Learning</i> Jeongwoo Park, Enrico Liscio and Pradeep Kumar Murukannaiah.....	654
<i>Prosody in Cascade and Direct Speech-to-Text Translation: a case study on Korean Wh-Phrases</i> Giulio Zhou, Tsz Kin Lam, Alexandra Birch and Barry Haddow .....	674
<i>Exploring the Potential of ChatGPT on Sentence Level Relations: A Focus on Temporal, Causal, and Discourse Relations</i> Chunkit Chan, Cheng Jiayang, Weiqi Wang, Yuxin Jiang, Tianqing Fang, Xin Liu and Yangqiu Song.....	684
<i>Backtracing: Retrieving the Cause of the Query</i> Rose E Wang, Pawan Wirawarn, Omar Khattab, Noah Goodman and Dorottya Demszky ....	722
<i>Unsupervised Multilingual Dense Retrieval via Generative Pseudo Labeling</i> Chao-Wei Huang, Chen-An Li, Tsu-Yuan Hsu, Chen-Yu Hsu and Yun-Nung Chen .....	736

<i>Investigating grammatical abstraction in language models using few-shot learning of novel noun gender</i> Priyanka Sukumaran, Conor Houghton and Nina Kazanina .....	747
<i>On-the-fly Denoising for Data Augmentation in Natural Language Understanding</i> Tianqing Fang, Wenxuan Zhou, Fangyu Liu, Hongming Zhang, Yangqiu Song and Muhao Chen	766
<i>Style Vectors for Steering Generative Large Language Models</i> Kai Konen, Sophie Jentsch, Diaoulé Diallo, Peer Schütt, Oliver Bensch, Roxanne El Baff, Dominik Opitz and Tobias Hecking .....	782
<i>Consistent Joint Decision-Making with Heterogeneous Learning Models</i> Hossein Rajaby Faghihi and Parisa Kordjamshidi .....	803
<i>Quantifying Association Capabilities of Large Language Models and Its Implications on Privacy Leakage</i> Hanyin Shao, Jie Huang, Shen Zheng and Kevin Chang .....	814
<i>Probing Critical Learning Dynamics of PLMs for Hate Speech Detection</i> Sarah Masud, Mohammad Aflah Khan, Vikram Goyal, Md Shad Akhtar and Tanmoy Chakraborty	826
<i>Emble: Reconstruction of Ancient Hebrew and Aramaic Texts Using Transformers</i> Niv Fono, Harel Moshayof, Eldar Karol, Itai Assraf and Mark Last .....	846
<i>Stateful Memory-Augmented Transformers for Efficient Dialogue Modeling</i> Qingyang Wu and Zhou Yu .....	853
<i>The Shape of Learning: Anisotropy and Intrinsic Dimensions in Transformer-Based Models</i> Anton Razzhigaev, Matvey Mikhailchuk, Elizaveta Goncharova, Ivan Oseledets, Denis Dimitrov and Andrey Kuznetsov .....	868
<i>MEDs for PETs: Multilingual Euphemism Disambiguation for Potentially Euphemistic Terms</i> Patrick Lee, Alain Chirino Trujillo, Diana Cuevas Plancarte, Olumide Ebenezer Ojo, Xinyi Liu, Iyanuoluwa Shode, Yuan Zhao, Anna Feldman and Jing Peng .....	875
<i>PromptExplainer: Explaining Language Models through Prompt-based Learning</i> Zijian Feng, Hanzhang Zhou, Zixiao Zhu and Kezhi Mao .....	882
<i>Do-Not-Answer: Evaluating Safeguards in LLMs</i> Yuxia Wang, Haonan Li, Xudong Han, Preslav Nakov and Timothy Baldwin .....	896
<i>Do Language Models Know When They're Hallucinating References?</i> Ayush Agrawal, Mirac Suzgun, Lester Mackey and Adam Tauman Kalai .....	912
<i>Bridging Cultural Nuances in Dialogue Agents through Cultural Value Surveys</i> Yong Cao, Min Chen and Daniel Hershcovich .....	929
<i>CEO: Corpus-based Open-Domain Event Ontology Induction</i> Nan Xu, Hongming Zhang and Jianshu Chen .....	946
<i>Rethinking STS and NLI in Large Language Models</i> Yuxia Wang, Minghan Wang and Preslav Nakov .....	965
<i>Learning High-Quality and General-Purpose Phrase Representations</i> Lihu Chen, Gael Varoquaux and Fabian M. Suchanek .....	983

<i>Explaining Language Model Predictions with High-Impact Concepts</i> Ruo Chen Zhao, Tan Wang, Yongjie Wang and Shafiq Joty . . . . .	995
<i>Understanding and Mitigating Spurious Correlations in Text Classification with Neighborhood Analysis</i> Oscar Chew, Hsuan-Tien Lin, Kai-Wei Chang and Kuan-Hao Huang . . . . .	1013
<i>On the Intractability to Synthesize Factual Inconsistencies in Summarization</i> Ge Luo, Weisi Fan, Miaoran Li, Youbiao He, Yinfei Yang and Forrest Sheng Bao . . . . .	1026
<i>IndiVec: An Exploration of Leveraging Large Language Models for Media Bias Detection with Fine-Grained Bias Indicators</i> Luyang Lin, Lingzhi Wang, Xiaoyan Zhao, Jing Li and Kam-Fai Wong . . . . .	1038
<i>Are Large Language Model-based Evaluators the Solution to Scaling Up Multilingual Evaluation?</i> Rishav Hada, Varun Gumma, Adrian de Wynter, Harshita Diddee, Mohamed Ahmed, Monojit Choudhury, Kalika Bali and Sunayana Sitaram . . . . .	1051
<i>Computational Morphology and Lexicography Modeling of Modern Standard Arabic Nominals</i> Christian Khairallah, Reham Marzouk, Salam Khalifa, Mayar Nassar and Nizar Habash . . . . .	1071
<i>Relabeling Minimal Training Subset to Flip a Prediction</i> Jinghan Yang, Linjie Xu and Lequan Yu . . . . .	1085
<i>Why Generate When You Can Discriminate? A Novel Technique for Text Classification using Language Models</i> Sachin Pawar, Nitin Ramrakhiani, Anubhav Sinha, Manoj Apte and Girish Keshav Palshikar . . . . .	1099
<i>Autism Detection in Speech – A Survey</i> Nadine Probol and Margot Mieskes . . . . .	1115
<i>Improving Multimodal Classification of Social Media Posts by Leveraging Image-Text Auxiliary Tasks</i> Danae Sanchez Villegas, Daniel Preotiuc-Pietro and Nikolaos Aletras . . . . .	1126
<i>What the Weight?! A Unified Framework for Zero-Shot Knowledge Composition</i> Carolin Holtermann, Markus Frohmann, Navid Rekasaz and Anne Lauscher . . . . .	1138
<i>IndiFoodVQA: Advancing Visual Question Answering and Reasoning with a Knowledge-Infused Synthetic Data Generation Pipeline</i> Pulkit Agarwal, Settaluri Lakshmi Sravanthi and Pushpak Bhattacharyya . . . . .	1158
<i>MAPLE: Micro Analysis of Pairwise Language Evolution for Few-Shot Claim Verification</i> Xia Zeng and Arkaitz Zubiaga . . . . .	1177
<i>Leveraging Open Information Extraction for More Robust Domain Transfer of Event Trigger Detection</i> David Dukić, Kiril Gashteovski, Goran Glavaš and Jan Snajder . . . . .	1197
<i>Exploring efficient zero-shot synthetic dataset generation for Information Retrieval</i> Tiago Almeida and Sérgio Matos . . . . .	1214
<i>Clustering-based Sampling for Few-Shot Cross-Domain Keyphrase Extraction</i> Prakmya Mishra, Lincy Pattanaik, Arunima Sundar, Nishant Yadav and Mayank Kulkarni . . . . .	1232
<i>Random Smooth-based Certified Defense against Text Adversarial Attack</i> Zeliang Zhang, Wei Yao, Susan Liang and Chenliang Xu . . . . .	1251

<i>Clarifying the Path to User Satisfaction: An Investigation into Clarification Usefulness</i> Hossein A. Rahmani, Xi Wang, Mohammad Aliannejadi, Mohammadmehdi Naghiaei and Emine Yilmaz.....	1266
<i>Efficiently Aligned Cross-Lingual Transfer Learning for Conversational Tasks using Prompt-Tuning</i> Lifu Tu, Jin Qu, Semih Yavuz, Shafiq Joty, Wenhao Liu, Caiming Xiong and Yingbo Zhou .	1278
<i>Correcting Language Model Outputs by Editing Salient Layers</i> Kshitij Mishra, Tamer Soliman, Anil Ramakrishna, Aram Galstyan and Anoop Kumar . . . . .	1295
<i>Improving Grounded Language Understanding in a Collaborative Environment by Interacting with Agents Through Help Feedback</i> Nikhil Mehta, Milagro Teruel, Xin Deng, Sergio Patricio Figueroa Sanz, Ahmed Hassan Awadal-lah and Julia Kiseleva . . . . .	1306
<i>Goodhart’s Law Applies to NLP’s Explanation Benchmarks</i> Jennifer Hsia, Danish Pruthi, Aarti Singh and Zachary Chase Lipton . . . . .	1322
<i>Syllable-level lyrics generation from melody exploiting character-level language model</i> Zhe Zhang, Karol Lasocki, Yi Yu and Atsuhiko Takasu . . . . .	1336
<i>Monolingual or Multilingual Instruction Tuning: Which Makes a Better Alpaca</i> Pinzhen Chen, Shaoxiong Ji, Nikolay Bogoychev, Andrey Kutuzov, Barry Haddow and Kenneth Heafield . . . . .	1347
<i>Prompt Perturbation Consistency Learning for Robust Language Models</i> Yao Qiang, Subhrangshu Nandi, Ninareh Mehrabi, Greg Ver Steeg, Anoop Kumar, Anna Rum-shisky and Aram Galstyan . . . . .	1357
<i>Enhancing Society-Undermining Disinformation Detection through Fine-Grained Sentiment Analysis Pre-Finetuning</i> Tsung-Hsuan Pan, Chung-Chi Chen, Hen-Hsen Huang and Hsin-Hsi Chen . . . . .	1371
<i>Minimal Distillation Schedule for Extreme Language Model Compression</i> Chen Zhang, Yang Yang, Qifan Wang, Jiahao Liu, Jingang Wang, Wei Wu and Dawei Song	1378
<i>Event Semantic Classification in Context</i> Haoyu Wang, Hongming Zhang, Kaiqiang Song, Dong Yu and Dan Roth . . . . .	1395
<i>Local and Global Contexts for Conversation</i> Zuoquan Lin and Xinyi Shen . . . . .	1408
<i>Aspect-based Key Point Analysis for Quantitative Summarization of Reviews</i> An Quang Tang, Xiuzhen Zhang and Minh Ngoc Dinh . . . . .	1419
<i>Improving Semantic Control in Discrete Latent Spaces with Transformer Quantized Variational Autoen-coders</i> Yingji Zhang, Danilo Carvalho, Marco Valentino, Ian Pratt-Hartmann and Andre Freitas . . .	1434
<i>High-quality Data-to-Text Generation for Severely Under-Resourced Languages with Out-of-the-box Large Language Models</i> Michela Lorandi and Anya Belz . . . . .	1451
<i>Antonym vs Synonym Distinction using InterlaCed Encoder NETworks (ICE-NET)</i> Muhammad Asif Ali, Yan HU, Jianbin Qin and Di Wang . . . . .	1462



<i>Predicting Machine Translation Performance on Low-Resource Languages: The Role of Domain Similarity</i>	Eric Khiu, Hasti Toossi, Jinyu Liu, Jiayu Li, David Anugraha, Juan Armando Parra Flores, Leandro Arcos Roman, A. Seza Doğruöz and En-Shiun Annie Lee . . . . .	1474
<i>Does CLIP Bind Concepts? Probing Compositionality in Large Image Models</i>	Martha Lewis, Nihal V. Nayak, Peilin Yu, Jack Merullo, Qinan Yu, Stephen Bach and Ellie Pavlick	1487
<i>Code-Switching and Back-Transliteration Using a Bilingual Model</i>	Daniel Weisberg Mitelman, Nachum Dershowitz and Kfir Bar . . . . .	1501
<i>Tsetlin Machine Embedding: Representing Words Using Logical Expressions</i>	Bimal Bhattarai, Ole-Christoffer Granmo, Lei Jiao, Rohan Kumar Yadav and Jivitesh Sharma	1512
<i>Reading Between the Tweets: Deciphering Ideological Stances of Interconnected Mixed-Ideology Communities</i>	Zihao He, Ashwin Rao, Siyi Guo, Negar Mokherian and Kristina Lerman . . . . .	1523
<i>Unified Embeddings for Multimodal Retrieval via Frozen LLMs</i>	Ziyang Wang, Heba Elfardy, Markus Dreyer, Kevin Small and Mohit Bansal . . . . .	1537
<i>Assessing the Portability of Parameter Matrices Trained by Parameter-Efficient Finetuning Methods</i>	Mohammed Sabry Mohammed and Anya Belz . . . . .	1548
<i>Exploiting Class Probabilities for Black-box Sentence-level Attacks</i>	Raha Moraffah and Huan Liu . . . . .	1557
<i>Learning Label Hierarchy with Supervised Contrastive Learning</i>	Ruixue Lian, William A. Sethares and Junjie Hu . . . . .	1569
<i>GrounDial: Human-norm Grounded Safe Dialog Response Generation</i>	Siwon Kim, Shuyang Dai, Mohammad Kachuee, Shayan Ray, Tara Taghavi and Sungroh Yoon	1582
<i>Trainable Hard Negative Examples in Contrastive Learning for Unsupervised Abstractive Summarization</i>	Haojie Zhuang, Wei Emma Zhang, Chang George Dong, Jian Yang and Quan Z. Sheng . . . . .	1589
<i>Low-Resource Counterspeech Generation for Indic Languages: The Case of Bengali and Hindi</i>	Mithun Das, Saurabh Kumar Pandey, Shivansh Sethi, Punyajoy Saha and Animesh Mukherjee	1601
<i>Teaching Probabilistic Logical Reasoning to Transformers</i>	Aliakbar Nafar, K. Brent Venable and Parisa Kordjamshidi . . . . .	1615
<i>On Measuring Context Utilization in Document-Level MT Systems</i>	Wafaa Mohammed and Vlad Niculae . . . . .	1633
<i>Solving NLP Problems through Human-System Collaboration: A Discussion-based Approach</i>	Masahiro Kaneko, Graham Neubig and Naoaki Okazaki . . . . .	1644
<i>Autoregressive Score Generation for Multi-trait Essay Scoring</i>	Heejin Do, Yunsu Kim and Gary Lee . . . . .	1659
<i>CMA-R: Causal Mediation Analysis for Explaining Rumour Detection</i>	Lin Tian, Xiuzhen Zhang and Jey Han Lau . . . . .	1667

<i>Morphology Aware Source Term Masking for Terminology-Constrained NMT</i> Ander Corral and Xabier Saralegi .....	1676
<i>Improving Backchannel Prediction Leveraging Sequential and Attentive Context Awareness</i> Yo-Han Park, Wencke Liermann, Yong-Seok Choi and Kong Joo Lee .....	1689
<i>SENSE-LM : A Synergy between a Language Model and Sensorimotor Representations for Auditory and Olfactory Information Extraction</i> Cédric Boscher, Christine Largeron, Véronique Eglin and Elöd Egyed-Zsigmond .....	1695
<i>Analyzing the Role of Part-of-Speech in Code-Switching: A Corpus-Based Study</i> Jie Chi and Peter Bell .....	1712
<i>In-Contextual Gender Bias Suppression for Large Language Models</i> Daisuke Oba, Masahiro Kaneko and Danushka Bollegala .....	1722
<i>Parameter-Efficient Fine-Tuning: Is There An Optimal Subset of Parameters to Tune?</i> Max Ploner and Alan Akbik .....	1743
<i>Contextualized Topic Coherence Metrics</i> Hamed Rahimi, David Mimno, Jacob Louis Hoover, Hubert Naacke, Camelia Constantin and Bernd Amann .....	1760
<i>PROMISE: A Proactive Multi-turn Dialogue Dataset for Information-seeking Intent Resolution</i> Yash Parag Butala, Siddhant Garg, Pratyay Banerjee and Amita Misra.....	1774
<i>CODET: A Benchmark for Contrastive Dialectal Evaluation of Machine Translation</i> Md Mahfuz Ibn Alam, Sina Ahmadi and Antonios Anastasopoulos .....	1790
<i>QAEVENT: Event Extraction as Question-Answer Pairs Generation</i> Milind Choudhary and Xinya Du .....	1860
<i>Sequence Shortening for Context-Aware Machine Translation</i> Pawel Maka, Yusuf Can Semerci, Jan Scholtes and Gerasimos Spanakis .....	1874
<i>Jigsaw Pieces of Meaning: Modeling Discourse Coherence with Informed Negative Sample Synthesis</i> Shubhankar Singh .....	1895
<i>Non-Exchangeable Conformal Language Generation with Nearest Neighbors</i> Dennis Thomas Ulmer, Chrysoula Zerva and Andre Martins .....	1909
<i>Evidentiality-aware Retrieval for Overcoming Abstractiveness in Open-Domain Question Answering</i> Yongho Song, Dahyun Lee, Myungha Jang, Seung-won Hwang, Kyungjae Lee, Dongha Lee and Jinyoung Yeo .....	1930
<i>Self-training Strategies for Sentiment Analysis: An Empirical Study</i> Haochen Liu, Sai Krishna Rallabandi, Yijing Wu, Parag Pravin Dakle and Preethi Raghavan	1944
<i>Language is All a Graph Needs</i> Ruosong Ye, Caiqi Zhang, Runhui Wang, Shuyuan Xu and Yongfeng Zhang .....	1955
<i>Unraveling the Dynamics of Semi-Supervised Hate Speech Detection: The Impact of Unlabeled Data Characteristics and Pseudo-Labeling Strategies</i> Florian Ludwig, Klara Dolos, Ana Alves-Pinto and Torsten Zesch .....	1974
<i>When do Generative Query and Document Expansions Fail? A Comprehensive Study Across Methods, Retrievers, and Datasets</i> Orion Weller, Kyle Lo, David Wadden, Dawn Lawrie, Benjamin Van Durme, Arman Cohan and Luca Soldaini.....	1987

<i>Can Large Language Models Understand Context?</i>	
Yilun Zhu, Joel Ruben Antony Moniz, Shruti Bhargava, Jiarui Lu, Dhivya Piraviperumal, Site Li, Yuan Zhang, Hong Yu and Bo-Hsiang Tseng . . . . .	2004
<i>Let's Negotiate! A Survey of Negotiation Dialogue Systems</i>	
Haolan Zhan, Yufei Wang, Zhuang Li, Tao Feng, Yuncheng Hua, Suraj Sharma, Lizhen Qu, Zhaleh Semnani Azad, Ingrid Zukerman and Reza Haf . . . . .	2019
<i>Towards Understanding Counseling Conversations: Domain Knowledge and Large Language Models</i>	
Younghun Lee, Dan Goldwasser and Laura Schwab Reese . . . . .	2032
<i>Better Explain Transformers by Illuminating Important Information</i>	
Linxin Song, Yan Cui, Ao Luo, Freddy Lecue and Irene Li . . . . .	2048
<i>Testing the Depth of ChatGPT's Comprehension via Cross-Modal Tasks Based on ASCII-Art: GPT3.5's Abilities in Regard to Recognizing and Generating ASCII-Art Are Not Totally Lacking</i>	
David Bayani . . . . .	2063
<i>Cross-lingual Editing in Multilingual Language Models</i>	
Himanshu Beniwal, Kowsik Nandagopan D and Mayank Singh . . . . .	2078
<i>Sorted LLaMA: Unlocking the Potential of Intermediate Layers of Large Language Models for Dynamic Inference</i>	
Parsa Kavehzadeh, Mojtaba Valipour, Marzieh S. Tahaei, Ali Ghodsi, Boxing Chen and Mehdi Rezagholizadeh . . . . .	2129
<i>AccentFold: A Journey through African Accents for Zero-Shot ASR Adaptation to Target Accents</i>	
Abraham Toluwase Owodunni, Aditya Yadavalli, Chris Chinenye Emezue, Tobi Olatunji and Clinton C Mbataku . . . . .	2146
<i>Hierarchical and Dynamic Prompt Compression for Efficient Zero-shot API Usage</i>	
Yichen Jiang, Marco Del Vecchio, Mohit Bansal and Anders Johannsen . . . . .	2162
<i>Fine-tuning CLIP Text Encoders with Two-step Paraphrasing</i>	
Hyunjae Kim, Seunghyun Yoon, Trung Bui, Handong Zhao, Quan Hung Tran, Franck Dernoncourt and Jaewoo Kang . . . . .	2175
<i>Generative Interpretation: Toward Human-Like Evaluation for Educational Question-Answer Pair Generation</i>	
Hyeonseok Moon, Jaewook Lee, Sugyeong Eo, Chanjun Park, Jaehyung Seo and Heuseok Lim . . . . .	2185
<i>Dive into the Chasm: Probing the Gap between In- and Cross-Topic Generalization</i>	
Andreas Waldis, Yufang Hou and Iryna Gurevych . . . . .	2197
<i>LLM-GEM: Large Language Model-Guided Prediction of People's Empathy Levels towards Newspaper Article</i>	
Md Rakibul Hasan, Md Zakir Hossain, Tom Gedeon and Shafin Rahman . . . . .	2215
<i>ICE-Score: Instructing Large Language Models to Evaluate Code</i>	
Terry Yue Zhuo . . . . .	2232
<i>CRese: Benchmark Data and Automatic Evaluation Framework for Recommending Eligibility Criteria from Clinical Trial Information</i>	
Siun Kim, Jung-Hyun Won, David Lee, Renqian Luo, Lijun Wu, Tao Qin and Howard Lee . . . . .	2243

<i>BMX: Boosting Natural Language Generation Metrics with Explainability</i> Christoph Leiter, Hoa Nguyen and Steffen Eger .....	2274
<i>Joint Inference of Retrieval and Generation for Passage Re-ranking</i> Wei Fang, Yung-Sung Chuang and James R. Glass .....	2289
<i>DialogStudio: Towards Richest and Most Diverse Unified Dataset Collection for Conversational AI</i> Jianguo Zhang, Kun Qian, Zhiwei Liu, Shelby Heinecke, Rui Meng, Ye Liu, Zhou Yu, Huan Wang, Silvio Savarese and Caiming Xiong .....	2299
<i>Exploring hybrid approaches to readability: experiments on the complementarity between linguistic features and transformers</i> Rodrigo Wilkens, Patrick Watrin, Rémi Cardon, Alice Pintard, Isabelle Gribomont and Thomas François .....	2316
<i>Establishing degrees of closeness between audio recordings along different dimensions using large-scale cross-lingual models</i> Maxime Fily, Guillaume Wisniewski, Severine Guillaume, Gilles Adda and Alexis Michaud	2332
<i>The Queen of England is not England’s Queen: On the Lack of Factual Coherency in PLMs</i> Paul Youssef, Jörg Schlötterer and Christin Seifert .....	2342
<i>HierarchyNet: Learning to Summarize Source Code with Heterogeneous Representations</i> Minh Huynh Nguyen, Nghi D. Q. Bui, Truong Son Hy, Long Tran-Thanh and Tien N Nguyen	2355
<i>Understanding the effects of language-specific class imbalance in multilingual fine-tuning</i> Vincent Jung and Lonneke Van Der Plas .....	2368
<i>NL2Formula: Generating Spreadsheet Formulas from Natural Language Queries</i> Wei Zhao, Zhitao Hou, Siyuan Wu, Yan Gao, Haoyu Dong, Yao Wan, Hongyu Zhang, Yulei Sui and Haidong Zhang .....	2377