

# **Findings of the Association for Computational Linguistics: NAACL 2024**

2024 Conference of the North American  
Chapter of the Association for  
Computational Linguistics

Mexico City, Mexico  
16-21 June 2024

Volume 1 of 6

ISBN: 978-1-7138-9899-3

**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>Structured Pruning for Large Language Models Using Coupled Components Elimination and Minor Fine-tuning</i>	
Honghe Zhang, Xiaolong Shi Xiaolong Shi, Jingwei Sun and Guangzhong Sun . . . . .	1
<i>Weight-Inherited Distillation for Task-Agnostic BERT Compression</i>	
Taiqiang Wu, Cheng Hou, Shanshan Lao, Jiayi Li, Ngai Wong, Zhe Zhao and Yujiu Yang . . . .	13
<i>Ignore Me But Don't Replace Me: Utilizing Non-Linguistic Elements for Pretraining on the Cybersecurity Domain</i>	
Eugene Jang, Jian Cui, Dayeon Yim, Youngjin Jin, Jin-Woo Chung, Seungwon Shin and Yongjae Lee . . . . .	29
<i>Extremely efficient online query encoding for dense retrieval</i>	
Nachshon Cohen, Yaron Fainstein and Guy Kushilevitz . . . . .	43
<i>DIVKNOWQA: Assessing the Reasoning Ability of LLMs via Open-Domain Question Answering over Knowledge Base and Text</i>	
Wenting Zhao, Ye Liu, Tong Niu, Yao Wan, Philip S. Yu, Shafiq Joty, Yingbo Zhou and Semih Yavuz . . . . .	51
<i>SpeedE: Euclidean Geometric Knowledge Graph Embedding Strikes Back</i>	
Aleksandar Pavlović and Emanuel Sallinger . . . . .	69
<i>Language Guided Exploration for RL Agents in Text Environments</i>	
Hitesh Golchha, Sahil Yerawar, Dhruvesh Patel, Soham Dan and Keerthiram Murugesan . . . . .	93
<i>GPT-who: An Information Density-based Machine-Generated Text Detector</i>	
Saranya Venkatraman, Adaku Uchendu and Dongwon Lee . . . . .	103
<i>DEED: Dynamic Early Exit on Decoder for Accelerating Encoder-Decoder Transformer Models</i>	
Peng Tang, Pengkai Zhu, Tian Li, Srikar Appalaraju, Vijay Mahadevan and R. Manmatha . . .	116
<i>Attention Alignment and Flexible Positional Embeddings Improve Transformer Length Extrapolation</i>	
Ta-Chung Chi, Ting-Han Fan and Alexander Rudnicky . . . . .	132
<i>Automatic Pair Construction for Contrastive Post-training</i>	
Canwen Xu, Corby Rosset, Ethan C. Chau, Luciano Del Corro, Shweti Mahajan, Julian McAuley, Jennifer Neville, Ahmed Hassan Awadallah and Nikhil Rao . . . . .	149
<i>Self-Checker: Plug-and-Play Modules for Fact-Checking with Large Language Models</i>	
Miaoran Li, Baolin Peng, Michel Galley, Jianfeng Gao and Zhu Zhang . . . . .	163
<i>Low-resource neural machine translation with morphological modeling</i>	
Antoine Nzeyimana . . . . .	182
<i>Self-Cleaning: Improving a Named Entity Recognizer Trained on Noisy Data with a Few Clean Instances</i>	
Zhendong Chu, Ruiyi Zhang, Tong Yu, Rajiv Jain, Vlad I Morariu, Jiuxiang Gu and Ani Nenkova	196
<i>VLUe: A New Benchmark and Multi-task Knowledge Transfer Learning for Vietnamese Natural Language Understanding</i>	
Phong Nguyen-Thuan Do, Son Quoc Tran, Phu Gia Hoang, Kiet Van Nguyen and Ngan Luu-Thuy Nguyen . . . . .	211

<i>LETI: Learning to Generate from Textual Interactions</i> Xingyao Wang, Hao Peng, Reyhaneh Jabbarvand and Heng Ji .....	223
<i>Bilateral Masking with prompt for Knowledge Graph Completion</i> yonghui kong, Cunhang Fan, Yujie Chen, Shuai Zhang, Zhao Lv and Jianhua Tao .....	240
<i>MiLe Loss: a New Loss for Mitigating the Bias of Learning Difficulties in Generative Language Models</i> Zhenpeng Su, Zijia Lin, Baixue Baixue, Hui Chen, Songlin Hu, Wei Zhou, Guiguang Ding and Xing W .....	250
<i>GOLD: Geometry Problem Solver with Natural Language Description</i> Jiaxin Zhang and Yashar Moshfeghi .....	263
<i>RoDia: A New Dataset for Romanian Dialect Identification from Speech</i> Rotaru Codruț, Nicolae Catalin Ristea and Radu Tudor Ionescu .....	279
<i>Examining Modularity in Multilingual LMs via Language-Specialized Subnetworks</i> Rochelle Choenni, Ekaterina Shutova and Dan Garrette .....	287
<i>Reverse Chain: A Generic-Rule for LLMs to Master Multi-API Planning</i> Yinger Zhang, Hui Cai, Xierui Song, Yicheng Chen, Rui Sun and Jing Zheng .....	302
<i>Incorporating Exponential Smoothing into MLP: a Simple but Effective Sequence Model</i> JiqunChu JiqunChu and Zuoquan Lin .....	326
<i>OpenFMNav: Towards Open-Set Zero-Shot Object Navigation via Vision-Language Foundation Models</i> Yuxuan Kuang, Hai Lin and Meng Jiang .....	338
<i>Comparing Two Model Designs for Clinical Note Generation; Is an LLM a Useful Evaluator of Consistency?</i> Nathan Brake and Thomas Schaaf .....	352
<i>VOLTA: Improving Generative Diversity by Variational Mutual Information Maximizing Autoencoder</i> Yueen Ma, DaFeng Chi, Jingjing Li, Kai Song, Yuzheng Zhuang and Irwin King .....	364
<i>EcoSpeak: Cost-Efficient Bias Mitigation for Partially Cross-Lingual Speaker Verification</i> Divya V Sharma .....	379
<i>Leveraging Contextual Information for Effective Entity Saliency Detection</i> Rajarshi Bhowmik, Marco Ponza, Atharva Tendle, Anant Gupta, Rebecca Jiang, Xingyu Lu, Qian Zhao and Daniel Preotiuc-Pietro .....	395
<i>LLM-as-a-Coauthor: Can Mixed Human-Written and Machine-Generated Text Be Detected?</i> Qihui Zhang, Chujie Gao, Dongping Chen, Yue Huang, Yixin Huang, Zhenyang Sun, Shilin Zhang, Weiye Li, Zhengyan Fu, Yao Wan and Lichao Sun .....	409
<i>A (More) Realistic Evaluation Setup for Generalisation of Community Models on Malicious Content Detection</i> Ivo Verhoeven, Pushkar Mishra, Rahel Beloch, Helen Yannakoudakis and Ekaterina Shutova	437
<i>Citation: A Key to Building Responsible and Accountable Large Language Models</i> Jie Huang and Kevin Chang .....	464
<i>Graph-Induced Syntactic-Semantic Spaces in Transformer-Based Variational AutoEncoders</i> Yingji Zhang, Marco Valentino, Danilo Carvalho, Ian Pratt-Hartmann and Andre Freitas .....	474

<i>Narrowing the Gap between Zero- and Few-shot Machine Translation by Matching Styles</i> Weiting Tan, Haoran Xu, Lingfeng Shen, Shuyue Stella Li, Kenton Murray, Philipp Koehn, Benjamin Van Durme and Yunmo Chen . . . . .	490
<i>Which Modality should I use - Text, Motif, or Image? : Understanding Graphs with Large Language Models</i> Debarati Das, Ishaan Gupta, Jaideep Srivastava and Dongyeop Kang . . . . .	503
<i>On-the-Fly Fusion of Large Language Models and Machine Translation</i> Hieu Hoang, Huda Khayrallah and Marcin Junczys-Dowmunt . . . . .	520
<i>READ: Improving Relation Extraction from an ADversarial Perspective</i> Dawei Li, William P Hogan and Jingbo Shang . . . . .	533
<i>REQUAL-LM: Reliability and Equity through Aggregation in Large Language Models</i> Sana Ebrahimi, Nima Shahbazi and Abolfazl Asudeh . . . . .	549
<i>Addressing Both Statistical and Causal Gender Fairness in NLP Models</i> Hannah Chen, Yangfeng Ji and David Evans . . . . .	561
<i>LLM-Rec: Personalized Recommendation via Prompting Large Language Models</i> Hanjia Lyu, Song Jiang, Hanqing Zeng, Yinglong Xia, Qifan Wang, Si Zhang, Ren Chen, Chris Leung, Jiajie Tang and Jiebo Luo . . . . .	583
<i>A Robust Semantics-based Watermark for Large Language Model against Paraphrasing</i> Jie Ren, Han Xu, Yiding Liu, Yingqian Cui, Shuaiqiang Wang, Dawei Yin and Jiliang Tang .	613
<i>Solving Data-centric Tasks using Large Language Models</i> Shraddha Barke, Christian Poelitz, Carina Suzana Negreanu, Benjamin Zorn, José Cambrone-ro, Andrew D. Gordon, Vu Le, Elnaz Nouri, Nadia Polikarpova, Advait Sarkar, Brian Slininger, Neil Toronto and Jack Williams . . . . .	626
<i>A Novel Paradigm Boosting Translation Capabilities of Large Language Models</i> Jiaxin Guo, Hao Yang, Zongyao Li, Daimeng Wei, Hengchao Shang and Xiaoyu Chen . . . . .	639
<i>Measuring Social Norms of Large Language Models</i> Ye Yuan, Kexin Tang, Jianhao Shen, Ming Zhang and Chenguang Wang . . . . .	650
<i>Source-Free Unsupervised Domain Adaptation for Question Answering via Prompt-Assisted Self-learning</i> Maxwell Juncheng Yin, Boyu Wang and Charles Ling . . . . .	700
<i>Hierarchical Attention Graph for Scientific Document Summarization in Global and Local Level</i> Chenlong Zhao, Xiwen Zhou, Xiaopeng Xie and Yong Zhang . . . . .	714
<i>LEETs-Dial: Linguistic Entrainment in End-to-End Task-oriented Dialogue systems</i> Nalin Kumar and Ondrej Dusek . . . . .	727
<i>Efficient Dependency Tree Sampling Without Replacement</i> Bogdan Mihai Dobre . . . . .	736
<i>Towards Better Generalization in Open-Domain Question Answering by Mitigating Context Memorization</i> Zixuan Zhang, Revanth Gangi Reddy, Kevin Small, Tong Zhang and Heng Ji . . . . .	742
<i>GEE! Grammar Error Explanation with Large Language Models</i> Yixiao Song, Kalpesh Krishna, Rajesh Bhatt, Kevin Gimpel and Mohit Iyyer . . . . .	754

<i>AdaRefiner: Refining Decisions of Language Models with Adaptive Feedback</i> Wanpeng Zhang and Zongqing Lu .....	782
<i>DivTOD: Unleashing the Power of LLMs for Diversifying Task-Oriented Dialogue Representations</i> Weihao Zeng, Dayuan Fu, Keqing He, Yejie Wang, Yukai Xu and Weiran Xu .....	800
<i>Teaching a Multilingual Large Language Model to Understand Multilingual Speech via Multi-Instructional Training</i> Pavel Denisov and Thang Vu .....	814
<i>CLEAN-EVAL: Clean Evaluation on Contaminated Large Language Models</i> Wenhong Zhu, Hongkun Hao, Zhiwei He, Yun-Ze Song, Jiao Yueyang, Yumeng Zhang, Hanxu Hu, Yiran Wei, Rui Wang and Hongyuan Lu .....	835
<i>R-BASS : Relevance-aided Block-wise Adaptation for Speech Summarization</i> Roshan Sharma, Ruchira Sharma, Hira Dhamyal, Rita Singh and Bhiksha Raj .....	848
<i>OVM, Outcome-supervised Value Models for Planning in Mathematical Reasoning</i> Fei Yu, Anningzhe Gao and Benyou Wang .....	858
<i>The Whole is Better than the Sum: Using Aggregated Demonstrations in In-Context Learning for Sequential Recommendation</i> Lei Wang and Ee-Peng Lim .....	876
<i>Bring Your Own KG: Self-Supervised Program Synthesis for Zero-Shot KGQA</i> Dhruv Agarwal, Rajarshi Das, Sopan Khosla and Rashmi Gangadharaiiah .....	896
<i>GraSAME: Injecting Token-Level Structural Information to Pretrained Language Models via Graph-guided Self-Attention Mechanism</i> Shuzhou Yuan and Michael Färber .....	920
<i>Can Public Large Language Models Help Private Cross-device Federated Learning?</i> Boxin Wang, Yibo Jacky Zhang, Yuan Cao, Bo Li, Hugh Brendan McMahan, Sewoong Oh, Zheng Xu and Manzil Zaheer .....	934
<i>LangNav: Language as a Perceptual Representation for Navigation</i> Bowen Pan, Rameswar Panda, SouYoung Jin, Rogerio Feris, Aude Oliva, Phillip Isola and Yoon Kim .....	950
<i>Planning and Editing What You Retrieve for Enhanced Tool Learning</i> Tenghao Huang, Dongwon Jung, Vaibhav Kumar, Mohammad Kachuee, Xiang Li, Puyang Xu and Muhao Chen .....	975
<i>Chart-based Reasoning: Transferring Capabilities from LLMs to VLMs</i> Victor Carbune, Hassan Mansoor, Fangyu Liu, Rahul Aralikkatte, Gilles Baechler, Jindong Chen and Abhanshu Sharma .....	989
<i>SLiM: Speculative Decoding with Hypothesis Reduction</i> Chi-Heng Lin, Shikhar Tuli, James Seale Smith, Yen-Chang Hsu, Yilin Shen and Hongxia Jin	1005
<i>REMATCH: Robust and Efficient Matching of Local Knowledge Graphs to Improve Structural and Semantic Similarity</i> Zohar Kachwala, Jisun An, Haewoon Kwak and Filippo Menczer .....	1018
<i>Modeling the Sacred: Considerations when Using Religious Texts in Natural Language Processing</i> Ben Hutchinson .....	1029

<i>Testing the Effect of Code Documentation on Large Language Model Code Understanding</i> William Macke and Michael Dean Doyle .....	1044
<i>Aligning Large Language Models with Recommendation Knowledge</i> Yuwei Cao, Nikhil Mehta, Xinyang Yi, Raghunandan Hulikal Keshavan, Lukasz Heldt, Lichan Hong, Ed H. Chi and Maheswaran Sathiamoorthy .....	1051
<i>OFA: A Framework of Initializing Unseen Subword Embeddings for Efficient Large-scale Multilingual Continued Pretraining</i> Yihong Liu, Peiqin Lin, Mingyang Wang and Hinrich Schuetze .....	1067
<i>SELF-EXPERTISE: Knowledge-based Instruction Dataset Augmentation for a Legal Expert Language Model</i> Minju Kim, Haein Jung and Myoung-Wan Koo .....	1098
<i>Re-evaluating the Need for Visual Signals in Unsupervised Grammar Induction</i> Boyi Li, Rodolfo Corona, Karttikeya Mangalam, Catherine Chen, Daniel Flaherty, Serge Belongie, Kilian Q Weinberger, Jitendra Malik, Trevor Darrell and Dan Klein .....	1113
<i>EDEntail: An Entailment-based Few-shot Text Classification with Extensional Definition</i> Zixiao Zhu, Junlang Qian, Zijian Feng, Hanzhang Zhou and Kezhi Mao .....	1124
<i>What Makes Math Word Problems Challenging for LLMs?</i> Kv Aditya Srivatsa and Ekaterina Kochmar .....	1138
<i>SMILE: Multimodal Dataset for Understanding Laughter in Video with Language Models</i> Lee Hyun, Kim Sung-Bin, Seungju Han, Youngjae Yu and Tae-Hyun Oh .....	1149
<i>T3M: Text Guided 3D Human Motion Synthesis from Speech</i> Wenshuo Peng, Kaipeng Zhang and Sai Qian Zhang .....	1168
<i>Deja vu: Contrastive Historical Modeling with Prefix-tuning for Temporal Knowledge Graph Reasoning</i> Miao Peng, Ben Liu, Wenjie Xu, Zihao Jiang, Jiahui Zhu and Min Peng .....	1178
<i>Explanation Extraction from Hierarchical Classification Frameworks for Long Legal Documents</i> Nishchal Prasad, Taoufiq Dkaki and Mohand Boughanem .....	1192
<i>Low-Rank Adaptation for Multilingual Summarization: An Empirical Study</i> Chenxi Whitehouse, Fantine Huot, Jasmijn Bastings, Mostafa Dehghani, Chu-Cheng Lin and Mirella Lapata .....	1202
<i>A Tree-of-Thoughts to Broaden Multi-step Reasoning across Languages</i> Leonardo Ranaldi, Giulia Pucci, Federico Ranaldi, Elena Sofia Ruzzetti and Fabio Massimo Zanzotto .....	1229
<i>Emergent Abilities in Reduced-Scale Generative Language Models</i> Sherin Muckatira, Vijeta Deshpande, Vladislav Lialin and Anna Rumshisky .....	1242
<i>Context Does Matter: Implications for Crowdsourced Evaluation Labels in Task-Oriented Dialogue Systems</i> Clemencia Siro, Mohammad Aliannejadi and Maarten de Rijke .....	1258
<i>Matching Varying-Length Texts via Topic-Informed and Decoupled Sentence Embeddings</i> Xixi Zhou, Chunbin Gu, Xin Jie, Jiajun Bu and Haishuai Wang .....	1274
<i>Instruction Tuning with Human Curriculum</i> Bruce W Lee, Hyunsoo Cho and Kang Min Yoo .....	1281

<i>Natural Language-based State Representation in Deep Reinforcement Learning</i> Md Masudur Rahman and Yexiang Xue .....	1310
<i>Learning Cross-Architecture Instruction Embeddings for Binary Code Analysis in Low-Resource Architectures</i> Junzhe Wang, Qiang Zeng and Lannan Luo .....	1320
<i>ReEval: Automatic Hallucination Evaluation for Retrieval-Augmented Large Language Models via Transferable Adversarial Attacks</i> Xiaodong Yu, Hao Cheng, Xiaodong Liu, Dan Roth and Jianfeng Gao .....	1333
<i>An Effective Automated Speaking Assessment Approach to Mitigating Data Scarcity and Imbalanced Distribution</i> Tien-Hong Lo, Fu-An Chao, Tzu-i WU, Yao-Ting Sung and Berlin Chen .....	1352
<i>GPT-Fathom: Benchmarking Large Language Models to Decipher the Evolutionary Path towards GPT-4 and Beyond</i> Shen Zheng, Yuyu Zhang, Yijie Zhu, Chenguang Xi, Pengyang Gao, Zhou Xun and Kevin Chang	1363
<i>Subword Attention and Post-Processing for Rare and Unknown Contextualized Embeddings</i> Raj Patel and Carlotta Domeniconi .....	1383
<i>UGIF-DataSet: A New Dataset for Cross-lingual, Cross-modal Sequential actions on the UI</i> Sagar Gubbi Venkatesh, Partha Talukdar and Srinu Narayanan .....	1390
<i>SimSCOOD: Systematic Analysis of Out-of-Distribution Generalization in Fine-tuned Source Code Models</i> Hossein Hajipour, Ning Yu, Cristian-Alexandru Staicu and Mario Fritz .....	1400
<i>Pruning as a Domain-specific LLM Extractor</i> Nan Zhang, Yanchi Liu, Xujiang Zhao, Wei Cheng, Runxue Bao, Rui Zhang, Prasenjit Mitra and Haifeng Chen .....	1417
<i>LLMRefine: Pinpointing and Refining Large Language Models via Fine-Grained Actionable Feedback</i> Wenda Xu, Daniel Deutsch, Mara Finkelstein, Juraj Juraska, Biao Zhang, Zhongtao Liu, William Yang Wang, Lei Li and Markus Freitag .....	1429
<i>Noisy Multi-Label Text Classification via Instance-Label Pair Correction</i> Pengyu Xu, Mingyang Song, Linkaida Liu, Bing Liu, Hongjian Sun, Liping Jing and Jian Yu	1446
<i>Composite Backdoor Attacks Against Large Language Models</i> Hai Huang, Zhengyu Zhao, Michael Backes, Yun Shen and Yang Zhang .....	1459
<i>Adapting Fake News Detection to the Era of Large Language Models</i> Jinyan Su, Claire Cardie and Preslav Nakov .....	1473
<i>MCAD: Multi-teacher Cross-modal Alignment Distillation for efficient image-text retrieval</i> Youbo Lei, Feifei He, Chen Chen, Yingbin Mo, Sijia Li, Defeng Xie and Haonan Lu .....	1491
<i>Large Language Models are Effective Text Rankers with Pairwise Ranking Prompting</i> Zhen Qin, Rolf Jagerman, Kai Hui, Honglei Zhuang, Junru Wu, Le Yan, Jiaming Shen, Tianqi Liu, Jialu Liu, Donald Metzler, Xuanhui Wang and Michael Bendersky .....	1504
<i>FedLFC: Towards Efficient Federated Multilingual Modeling with LoRA-based Language Family Clustering</i> Zhihan Guo, Yifei Zhang, Zhuo Zhang, Zenglin Xu and Irwin King .....	1519



<i>Gaussian Process Optimization for Adaptable Multi-Objective Text Generation using Linearly-Weighted Language Models</i>	
Mohammad Mahdi Abdollah Pour, Ali Pesaranghader, Eldan Cohen and Scott Sanner . . . . .	1529
<i>Groundedness in Retrieval-augmented Long-form Generation: An Empirical Study</i>	
Alessandro Stolfo . . . . .	1537
<i>TagDebias: Entity and Concept Tagging for Social Bias Mitigation in Pretrained Language Models</i>	
Mehrmaz Moslemi and Amal Zouaq . . . . .	1553
<i>Improving Absent Keyphrase Generation with Diversity Heads</i>	
Edwin Thomas and Sowmya Vajjala . . . . .	1568
<i>mOthello: When Do Cross-Lingual Representation Alignment and Cross-Lingual Transfer Emerge in Multilingual Models?</i>	
Tianze Hua, Tian Yun and Ellie Pavlick . . . . .	1585
<i>Discovering and Mitigating Indirect Bias in Attention-Based Model Explanations</i>	
Farsheed Haque, Depeng Xu and Shuhan Yuan . . . . .	1599
<i>i-Code V2: An Autoregressive Generation Framework over Vision, Language, and Speech Data</i>	
Ziyi Yang, Mahmoud Khademi, Yichong Xu, Reid Pryzant, Yuwei Fang, Chenguang Zhu, Dongdong Chen, Yao Qian, Xuemei Gao, Yi-Ling Chen, Robert Gmyr, Naoyuki Kanda, Noel C Codella, Bin Xiao, Yu Shi, Lu Yuan, Takuya Yoshioka, Michael Zeng and Xuedong Huang . . . . .	1615
<i>Think While You Write: Hypothesis Verification Promotes Faithful Knowledge-to-Text Generation</i>	
Yifu Qiu, Varun R. Embar, Shay B Cohen and Benjamin Han . . . . .	1628
<i>It's All Relative! – A Synthetic Query Generation Approach for Improving Zero-Shot Relevance Prediction</i>	
Aditi Chaudhary, Karthik Raman and Michael Bendersky . . . . .	1645
<i>RS-DPO: A Hybrid Rejection Sampling and Direct Preference Optimization Method for Alignment of Large Language Models</i>	
Saeed Khaki, JinJin Li, Lan Ma, Liu Yang and Prathap Ramachandra . . . . .	1665
<i>Hypernetwork-Assisted Parameter-Efficient Fine-Tuning with Meta-Knowledge Distillation for Domain Knowledge Disentanglement</i>	
Changqun Li, Linlin Wang, Xin Alex Lin, Shizhou Huang and Liang He . . . . .	1681
<i>MICo: Preventative Detoxification of Large Language Models through Inhibition Control</i>	
Roy Siegelmann, Ninareh Mehrabi, Palash Goyal, Prasoon Goyal, Lisa Bauer, Jwala Dhamala, Aram Galstyan, Rahul Gupta and Reza Ghanadan . . . . .	1696
<i>Reinforcement Learning with Token-level Feedback for Controllable Text Generation</i>	
Wendi Li, Wei Wei, Kaihe xu, Wenfeng xie, Dangyang Chen and Yu Cheng . . . . .	1704
<i>CoMM: Collaborative Multi-Agent, Multi-Reasoning-Path Prompting for Complex Problem Solving</i>	
Pei Chen, Shuai Zhang and Boran Han . . . . .	1720
<i>Tokenization Matters: Navigating Data-Scarce Tokenization for Gender Inclusive Language Technologies</i>	
Anaelia Ovalle, Ninareh Mehrabi, Palash Goyal, Jwala Dhamala, Kai-Wei Chang, Richard Zemel, Aram Galstyan, Yuval Pinter and Rahul Gupta . . . . .	1739
<i>AdaPT: A Set of Guidelines for Hyperbolic Multimodal Multilingual NLP</i>	
Ramit Sawhney, Shrey Pandit, Vishwa Shah, Megh Thakkar and Shafiq Joty . . . . .	1757

<i>More Samples or More Prompts? Exploring Effective Few-Shot In-Context Learning for LLMs with In-Context Sampling</i>	
Bingsheng Yao, Guiming Hardy Chen, Ruishi Zou, Yuxuan Lu, Jiachen Li, Shao Zhang, Yisi Sang, Sijia Liu, James Hendler and Dakuo Wang . . . . .	1772
<i>ZSEE: A Dataset based on Zeolite Synthesis Event Extraction for Automated Synthesis Platform</i>	
Song He, Xin Peng, Yihan Cai, Xin Li, Zhiqing Yuan, WenLi Du and Weimin Yang . . . . .	1791
<i>Mitigating Hallucination in Abstractive Summarization with Domain-Conditional Mutual Information</i>	
Kyubung Chae, Jaepill choi, Yohan Jo and Taesup Kim . . . . .	1809
<i>Adversarial DPO: Harnessing Harmful Data for Reducing Toxicity with Minimal Impact on Coherence and Evasiveness in Dialogue Agents</i>	
San Kim and Gary Lee . . . . .	1821
<i>Prompt Space Optimizing Few-shot Reasoning Success with Large Language Models</i>	
Fobo Shi, Peijun Qing, Dong Yang, Nan Wang, Youbo Lei, Haonan Lu, Xiaodong Lin and Duan-tengchuan Li . . . . .	1836
<i>DAGCN: Distance-based and Aspect-oriented Graph Convolutional Network for Aspect-based Sentiment Analysis</i>	
Zhihao Wang, Bo Zhang, Ru Yang, Chang Guo and Maozhen Li . . . . .	1863
<i>Connecting the Dots: Inferring Patent Phrase Similarity with Retrieved Phrase Graphs</i>	
Zhuoyi Peng and Yi Yang . . . . .	1877
<i>Self-Regulated Sample Diversity in Large Language Models</i>	
Mingyue Liu, Jonathan Frawley, Sarah Wyrer, Hubert P. H. Shum, Sara L. Uckelman, Sue Black and Chris G. Willcocks . . . . .	1891
<i>Methods, Applications, and Directions of Learning-to-Rank in NLP Research</i>	
Justin Lee, Gabriel Bernier-Colborne, Tegan Maharaj and Sowmya Vajjala . . . . .	1900
<i>When Quantization Affects Confidence of Large Language Models?</i>	
Irina Proskurina, Luc Brun, Guillaume Metzler and Julien Velcin . . . . .	1918
<i>MedCycle: Unpaired Medical Report Generation via Cycle-Consistency</i>	
Elad Hirsch, Gefen Dawidowicz and Ayellet Tal . . . . .	1929
<i>Beta-LR: Interpretable Logical Reasoning based on Beta Distribution</i>	
Yizhuo Ma, Ke Qin and Shuang Liang . . . . .	1945
<i>Applications of BERT Models Towards Automation of Clinical Coding in Icelandic</i>	
Haraldur Orri Hauksson and Hafsteinn Einarsson . . . . .	1956
<i>Tell me who you are and I tell you how you argue": Predicting Stances and Arguments for Stakeholder Groups</i>	
Philipp Heinisch, Lorik Dumani, Philipp Cimiano and Ralf Schenkel . . . . .	1968
<i>Psychometric Predictive Power of Large Language Models</i>	
Tatsuki Kuribayashi, Yohei Oseki and Timothy Baldwin . . . . .	1983
<i>Large Language Models Sensitivity to The Order of Options in Multiple-Choice Questions</i>	
Pouya Pezeshkpour and Estevam Hruschka . . . . .	2006
<i>PEEB: Part-based Image Classifiers with an Explainable and Editable Language Bottleneck</i>	
Thang M. Pham, Peijie Chen, Tin Nguyen, Seunghyun Yoon, Trung Bui and Anh Totti Nguyen	

<i>Ethos: Rectifying Language Models in Orthogonal Parameter Space</i> Lei Gao, Yue Niu, Tingting Tang, Salman Avestimehr and Murali Annavaram . . . . .	2054
<i>Crafting In-context Examples according to LMs' Parametric Knowledge</i> Yoonsang Lee, Pranav Atreya, Xi Ye and Eunsol Choi . . . . .	2069
<i>ICXML: An In-Context Learning Framework for Zero-Shot Extreme Multi-Label Classification</i> Yaxin Zhu and Hamed Zamani . . . . .	2086
<i>CLGSI: A Multimodal Sentiment Analysis Framework based on Contrastive Learning Guided by Sentiment Intensity</i> Yang Yang, Xunde Dong and Yupeng Qiang . . . . .	2099
<i>Interpreting Answers to Yes-No Questions in Dialogues from Multiple Domains</i> Zijie Wang, Farzana Rashid and Eduardo Blanco . . . . .	2111
<i>Enhancing Perception: Refining Explanations of News Claims with LLM Conversations</i> Yi-Li Hsu, Jui-Ning Chen, Yang Fan Chiang, Shang-Chien Liu, Aiping Xiong and Lun-Wei Ku 2129	
<i>How Interpretable are Reasoning Explanations from Prompting Large Language Models?</i> Yeo Wei Jie, Ranjan Satapathy, Rick Siow Mong Goh and Erik Cambria . . . . .	2148
<i>Plug-in Language Model: Controlling Text Generation with a Simple Regression Model</i> Nai-Chi Yang, Wei-Yun Ma and Pu-Jen Cheng . . . . .	2165
<i>Signer Diversity-driven Data Augmentation for Signer-Independent Sign Language Translation</i> honghaofu honghaofu, Liang Zhang, Biao Fu, Rui Zhao, Jinsong Su, Xiaodong Shi and Yidong Chen . . . . .	2182
<i>A Systematic Analysis of Subwords and Cross-Lingual Transfer in Multilingual Translation</i> Francois Meyer and Jan Buys . . . . .	2194
<i>Multi-Granularity Guided Fusion-in-Decoder</i> Eunseong Choi, Hyeri Lee and Jongwuk Lee . . . . .	2201
<i>Group Fairness in Multilingual Speech Recognition Models</i> Anna Katrine van Zee, Marc van Zee and Anders Søgaard . . . . .	2213
<i>Rethinking Machine Ethics – Can LLMs Perform Moral Reasoning through the Lens of Moral Theories?</i> Jingyan Zhou, Minda Hu, Junan Li, Xiaoying Zhang, Xixin Wu, Irwin King and Helen M. Meng 2227	
<i>Role Prompting Guided Domain Adaptation with General Capability Preserve for Large Language Models</i> Rui Wang, Fei Mi, Yi Chen, Boyang Xue, Hongru Wang, Qi Zhu, Kam-Fai Wong and Ruifeng Xu . . . . .	2243
<i>BERTweet's TACO Fiesta: Contrasting Flavors On The Path Of Inference And Information-Driven Argument Mining On Twitter</i> Marc Feger and Stefan Dietze . . . . .	2256
<i>Testing the limits of logical reasoning in neural and hybrid models</i> Manuel Vargas Guzman, Jakub Szymanik and Maciej Malicki . . . . .	2267
<i>METAL: Towards Multilingual Meta-Evaluation</i> Rishav Hada, Varun Gumma, Mohamed Ahmed, Kalika Bali and Sunayana Sitaram . . . . .	2280

<i>AGIEval: A Human-Centric Benchmark for Evaluating Foundation Models</i>	
Wanjun Zhong, Ruixiang Cui, Yiduo Guo, Yaobo Liang, Shuai Lu, Yanlin Wang, Amin Saied, Weizhu Chen and Nan Duan .....	2299
<i>Product Description and QA Assisted Self-Supervised Opinion Summarization</i>	
Tejpal Singh Siledar, Rupasai Rangaraju, Sankara Sri Raghava Ravindra Muddu, Suman Banerjee, Amey Patil, Sudhanshu Shekhar Singh, Muthusamy Chelliah, Nikesh Garera, Swaprava Nath and Pushpak Bhattacharyya .....	2315
<i>COMEM: In-Context Retrieval-Augmented Mass-Editing Memory in Large Language Models</i>	
Shanbao Qiao, Xuebing Liu and Seung-Hoon Na .....	2333
<i>Content-Specific Humorous Image Captioning Using Incongruity Resolution Chain-of-Thought</i>	
Kohtaro Tanaka, Kohei Uehara, Lin Gu, Yusuke Mukuta and Tatsuya Harada .....	2348
<i>Denoising Attention for Query-aware User Modeling</i>	
Elias Bassani, Pranav Kasela and Gabriella Pasi .....	2368
<i>A Lightweight Mixture-of-Experts Neural Machine Translation Model with Stage-wise Training Strategy</i>	
Fan Zhang, Mei Tu, Song Liu and Jinyao Yan .....	2381
<i>BEAR: A Unified Framework for Evaluating Relational Knowledge in Causal and Masked Language Models</i>	
Jacek Wiland, Max Ploner and Alan Akbik .....	2393
<i>Conformal Intent Classification and Clarification for Fast and Accurate Intent Recognition</i>	
Floris den Hengst, Ralf Wolter, Patrick Altmeyer and Arda Kaygan .....	2412
<i>Anonymity at Risk? Assessing Re-Identification Capabilities of Large Language Models in Court Decisions</i>	
Alex Nyffenegger, Matthias Stürmer and Joel Niklaus .....	2433
<i>X-LLaVA: Optimizing Bilingual Large Vision-Language Alignment</i>	
DongJae Shin, HyeonSeok Lim, Inho Won, ChangSu Choi, Minjun Kim, SeungWoo Song, HanGyeol Yoo, SangMin Kim and KyungTae Lim .....	2463
<i>Why So Gullible? Enhancing the Robustness of Retrieval-Augmented Models against Counterfactual Noise</i>	
Giwon Hong, Jeonghwan Kim, Junmo Kang, Sung-Hyon Myaeng and Joyce Jiyoung Whang	2474
<i>Heterogeneity over Homogeneity: Investigating Multilingual Speech Pre-Trained Models for Detecting Audio Deepfake</i>	
Orchid Chetia Phukan, Gautam Siddharth Kashyap, Arun Balaji Buduru and Rajesh Sharma	2496
<i>Identifying Self-Disclosures of Use, Misuse and Addiction in Community-based Social Media Posts</i>	
Chenghao Yang, Tuhin Chakrabarty, Karli R Hochstatter, Melissa N Slavin, Nabila El-Bassel and Smaranda Muresan .....	2507
<i>Self-Adaptive Sampling for Accurate Video Question Answering on Image Text Models</i>	
Wei Han, Hui Chen, Min-Yen Kan and Soujanya Poria .....	2522
<i>Towards an On-device Agent for Text Rewriting</i>	
Yun Zhu, Yinxiao Liu, Felix Stahlberg, Shankar Kumar, Yu-Hui Chen, Liangchen Luo, Lei Shu, Renjie Liu, Jindong Chen and Lei Meng .....	2535
<i>Tailoring Vaccine Messaging with Common-Ground Opinions</i>	
Rickard Stureborg, Sanxing Chen, Roy Xie, Aayushi Kunjal Patel, Christopher Li, Chloe Zhu, Tingnan Hu, Jun Yang and Bhuwan Dhingra .....	2553

<i>Best of Both Worlds: A Pliable and Generalizable Neuro-Symbolic Approach for Relation Classification</i> Robert Vacareanu, Fahmida Alam, Md Asiful Islam, Haris Riaz and Mihai Surdeanu . . . . .	2576
<i>Q-Tuning: Queue-based Prompt Tuning for Lifelong Few-shot Language Learning</i> Yanhui Guo, Shaoyuan Xu, Jinmiao Fu, Jia Liu, Chaosheng Dong and Bryan Wang . . . . .	2595
<i>In-Context Example Ordering Guided by Label Distributions</i> Zhichao Xu, Daniel Cohen, Bei Wang and Vivek Srikumar . . . . .	2623
<i>Beyond Surface Similarity: Detecting Subtle Semantic Shifts in Financial Narratives</i> Jiaxin Liu, Yi Yang and Kar Yan Tam . . . . .	2641
<i>Laying Anchors: Semantically Priming Numerals in Language Modeling</i> Mandar Sharma, Rutuja Taware, Pravesh Koirala, Nikhil Muralidhar and Naren Ramakrishnan	2653
<i>UEGP: Unified Expert-Guided Pre-training for Knowledge Rekindle</i> Yutao Mou, Kexiang Wang, Jianhe Lin, Dehong Ma, Jun Fan, Daiting Shi, Zhicong Cheng, Gu Simiu, Dawei Yin and Weiran Xu . . . . .	2661
<i>LatticeGen: Hiding Generated Text in a Lattice for Privacy-Aware Large Language Model Generation on Cloud</i> Mengke Zhang, Tianxing He, Tianle Wang, Lu Mi, Niloofar Mireshghallah, Binyi Chen, Hao Wang and Yulia Tsvetkov . . . . .	2674
<i>HateModerate: Testing Hate Speech Detectors against Content Moderation Policies</i> Jiangrui Zheng, Xueqing Liu, Mirazul Haque, Xing Qian, Guanqun Yang and Wei Yang . . .	2691
<i>Compensate Quantization Errors: Make Weights Hierarchical to Compensate Each Other</i> Yifei Gao, Jie Ou, Lei Wang, Yuting Xiao, xiangzhiyuan xiangzhiyuan, Ruiting Dai and Jun Cheng . . . . .	2711
<i>Contrastive Preference Learning for Neural Machine Translation</i> Jianfei He, Shichao Sun, Sen Peng, Jie Xu, Xiaohua Jia and Wenjie Li . . . . .	2723
<i>SocREval: Large Language Models with the Socratic Method for Reference-free Reasoning Evaluation</i> Hangfeng He, Hongming Zhang and Dan Roth . . . . .	2736
<i>Multilingual Machine Translation with Large Language Models: Empirical Results and Analysis</i> Wenhao Zhu, Hongyi Liu, Qingxiu Dong, Jingjing Xu, Shujian Huang, Lingpeng Kong, Jiajun Chen and Lei Li . . . . .	2765
<i>Unleashing the Power of LLMs in Court View Generation by Stimulating Internal Knowledge and Incorporating External Knowledge</i> Yifei Liu, Yiquan Wu, Ang Li, Yating Zhang, Changlong Sun, Weiming Lu, Fei Wu and Kun Kuang . . . . .	2782
<i>Prompting Vision-Language Models For Aspect-Controlled Generation of Referring Expressions</i> Danfeng Guo, Sanchit Agarwal, Arpit Gupta, Jiun-Yu Kao, Emre Barut, Tagyoung Chung, Jing Huang and Mohit Bansal . . . . .	2793
<i>Task-Agnostic Detector for Insertion-Based Backdoor Attacks</i> Weimin Lyu, Xiao Lin, Songzhu Zheng, Lu Pang, Haibin Ling, Susmit Jha and Chao Chen .	2808
<i>Uncertainty Estimation on Sequential Labeling via Uncertainty Transmission</i> Jianfeng He, Linlin Yu, Shuo Lei, Chang-Tien Lu and Feng Chen . . . . .	2823

<i>Exploring Language Model’s Code Generation Ability with Auxiliary Functions</i> Seonghyeon Lee, Sanghwan Jang, Seongbo Jang, Dongha Lee and Hwanjo Yu . . . . .	2836
<i>Crossing Linguistic Horizons: Finetuning and Comprehensive Evaluation of Vietnamese Large Language Models</i> Sang T. Truong, Duc Quang Nguyen, Toan Nguyen, Dong Le, Nhi Ngoc Truong, Tho Quan and Sanmi Koyejo . . . . .	2849
<i>GoT: Effective Graph-of-Thought Reasoning in Language Models</i> Yao Yao, Zuchao Li and hai zhao . . . . .	2901
<i>Enhancing the General Agent Capabilities of Low-Paramter LLMs through Tuning and Multi-Branch Reasoning</i> Qinhao Zhou, Zihan Zhang, Xiang Xiang, Ke Wang, Yuchuan Wu and Yongbin Li . . . . .	2922
<i>MuMath: Multi-perspective Data Augmentation for Mathematical Reasoning in Large Language Models</i> Weihao You, Shuo Yin, Xudong Zhao, Zhilong Ji, Guoqiang Zhong and Jinfeng Bai . . . . .	2932
<i>Tram: A Token-level Retrieval-augmented Mechanism for Source Code Summarization</i> Tong Ye, Lingfei Wu, Tengfei Ma, Xuhong Zhang, Yangkai Du, Peiyu Liu, Shouling Ji and Wenhai Wang . . . . .	2959
<i>UNO-DST: Leveraging Unlabelled Data in Zero-Shot Dialogue State Tracking</i> Chuang Li, Yan Zhang, Min-Yen Kan and Haizhou Li . . . . .	2972
<i>Evaluating Step-by-Step Reasoning through Symbolic Verification</i> YiFan Zhang, Hanlin Zhang, Li Erran Li and Eric P. Xing . . . . .	2984
<i>Multi-Review Fusion-in-Context</i> Aviv Slobodkin, Ori Shapira, Ran Levy and Ido Dagan . . . . .	3003
<i>Retrieving Examples from Memory for Retrieval Augmented Neural Machine Translation: A Systematic Comparison</i> Maxime Bouthors, Josep Crego and François Yvon . . . . .	3022
<i>Extending Input Contexts of Language Models through Training on Segmented Sequences</i> Petros Karypis, Julian McAuley and George Karypis . . . . .	3040
<i>Reason from Fallacy: Enhancing Large Language Models’ Logical Reasoning through Logical Fallacy Understanding</i> Yanda Li, Dixuan Wang, Jiaqing Liang, Guochao Jiang, Qianyu He, Yanghua Xiao and Deqing Yang . . . . .	3053
<i>Exploring Automated Distractor Generation for Math Multiple-choice Questions via Large Language Models</i> Wanyong Feng, Jaewook Lee, Hunter McNichols, Alexander Scarlatos, Digory Smith, Simon Woodhead, Nancy Otero Ornelas and Andrew Lan . . . . .	3067
<i>Aspect-based Sentiment Analysis with Context Denoising</i> Yuanhe Tian, Chang Liu, Yan Song, Fei Xia and Yongdong Zhang . . . . .	3083
<i>IruMozhi: Automatically classifying diglossia in Tamil</i> Kabilan Prasanna and Aryaman Arora . . . . .	3096

<i>RENOVI: A Benchmark Towards Remediating Norm Violations in Socio-Cultural Conversations</i>	
Haolan Zhan, Zhuang Li, Xiaoxi Kang, Tao Feng, Yuncheng Hua, Lizhen Qu, Yi Ying , Mei Rianto Chandra, Kelly Rosalin, Jureynolds Jureynolds, Suraj Sharma, Shilin QU, Linhao Luo, Ingrid Zukerman, Lay-Ki Soon, Zhaleh Semnani Azad and Reza Haf . . . . .	3104
<i>Human-in-the-Loop Synthetic Text Data Inspection with Provenance Tracking</i>	
Hong Jin Kang, Fabrice Y Harel-Canada, Muhammad Ali Gulzar, Nanyun Peng and Miryung Kim . . . . .	3118
<i>COMMIT: Code-Mixing English-Centric Large Language Model for Multilingual Instruction Tuning</i>	
Jaeseong Lee, YeonJoon Jung and seung-won hwang . . . . .	3130
<i>DiLM: Distilling Dataset into Language Model for Text-level Dataset Distillation</i>	
Aru Maekawa, Satoshi Kosugi, Kotaro Funakoshi and Manabu Okumura . . . . .	3138
<i>MindAgent: Emergent Gaming Interaction</i>	
Ran Gong, Qiuyuan Huang, Xiaojian Ma, Yusuke Noda, Zane Durante, Zilong Zheng, Demetri Terzopoulos, Li Fei-Fei, Jianfeng Gao and Hoi Vo . . . . .	3154
<i>BotChat: Evaluating LLMs’ Capabilities of Having Multi-Turn Dialogues</i>	
Haodong Duan, Jueqi Wei, Chonghua Wang, Hongwei Liu, Yixiao Fang, Songyang Zhang, Dahua Lin and Kai Chen . . . . .	3184
<i>Learning Mutually Informed Representations for Characters and Subwords</i>	
Yilin Wang, Xinyi Hu and Matthew R. Gormley . . . . .	3201
<i>A Novel Two-step Fine-tuning Framework for Transfer Learning in Low-Resource Neural Machine Translation</i>	
Yuan Gao, Feng Hou and Ruili Wang . . . . .	3214
<i>Enhancing Cross-lingual Sentence Embedding for Low-resource Languages with Word Alignment</i>	
Zhongtao Miao, Qiyu Wu, Kaiyan Zhao, Zilong Wu and Yoshimasa Tsuruoka . . . . .	3225
<i>C<sup>3</sup>LPGCN: Integrating Contrastive Learning and Cooperative Learning with Prompt into Graph Convolutional Network for Aspect-based Sentiment Analysis</i>	
Ye He, Shihao Zou, YuzheChen YuzheChen and Xianying Huang . . . . .	3237
<i>Visual Enhanced Entity-Level Interaction Network for Multimodal Summarization</i>	
Haolong Yan, Binghao Tang, Boda Lin, Gang Zhao and Si Li . . . . .	3248
<i>Knowledgeable In-Context Tuning: Exploring and Exploiting Factual Knowledge for In-Context Learning</i>	
Jianing Wang, Chengyu Wang, Chuanqi Tan, Jun Huang and Ming Gao . . . . .	3261
<i>Time Machine GPT</i>	
Felix Drinkall, Eghbal Rahimikia, Janet B. Pierrehumbert and Stefan Zohren . . . . .	3281
<i>An End-to-End Submodular Framework for Data-Efficient In-Context Learning</i>	
Lilly Kumari, Shengjie Wang, Arnav Mohanty Das, Tianyi Zhou and Jeff Bilmes . . . . .	3293
<i>Teaching Llama a New Language Through Cross-Lingual Knowledge Transfer</i>	
Hele-Andra Kuulmets, Taido Purason, Agnes Luhtaru and Mark Fishel . . . . .	3309
<i>Simulating Opinion Dynamics with Networks of LLM-based Agents</i>	
Yun-Shiuan Chuang, Agam Goyal, Nikunj Harlalka, Siddharth Suresh, Robert D. Hawkins, Sijia Yang, Dhavan V. Shah, Junjie Hu and Timothy T. Rogers . . . . .	3326

<i>Probing the Category of Verbal Aspect in Transformer Language Models</i> Anisia Katinskaia and Roman Yangarber .....	3347
<i>A Measure for Transparent Comparison of Linguistic Diversity in Multilingual NLP Data Sets</i> Tanja Samardzic, Ximena Gutierrez, Christian Bentz, Steven Moran and Olga Pelloni .....	3367
<i>Beyond Read-Only: Crafting a Comprehensive Chinese Text-to-SQL Dataset for Database Manipulation and Query</i> Xi Chen, Jinguo You, Likun Likun and Xiang Li .....	3383
<i>Normalizing without Modernizing: Keeping Historical Wordforms of Middle French while Reducing Spelling Variants</i> Raphael Rubino, Johanna Gerlach, Jonathan David Mutal and Pierrette Bouillon .....	3394
<i>Anti-LM Decoding for Zero-shot In-context Machine Translation</i> Suzanna Sia, Alexandra DeLucia and Kevin Duh .....	3403
<i>Defending Against Weight-Poisoning Backdoor Attacks for Parameter-Efficient Fine-Tuning</i> Shuai Zhao, Leilei Gan, Anh Tuan Luu, Jie Fu, Lingjuan Lyu, Meihuizi Jia and Jinming Wen	3421
<i>Select and Summarize: Scene Saliency for Movie Script Summarization</i> Rohit Saxena and Frank Keller .....	3439
<i>Don't be a Fool: Pooling Strategies in Offensive Language Detection from User-Intended Adversarial Attacks</i> Seunguk Yu, Juhwan Choi and YoungBin Kim .....	3456
<i>Z-GMOT: Zero-shot Generic Multiple Object Tracking</i> Kim Hoang Tran, Anh Duy Le Dinh, Tien-Phat Nguyen, Thinh Phan, Pha Nguyen, Khoa Luu, Donald Adjeroh, Gianfranco Doretto and Ngan Hoang Le .....	3468
<i>NLP for Counterspeech against Hate: A Survey and How-To Guide</i> Helena Bonaldi, Yi-Ling Chung, Gavin Abercrombie and Marco Guerini .....	3480
<i>PRODIGy: a PROFILE-based Dialogue Generation dataset</i> Daniela Occhipinti, Serra Sinem Tekiroglu and Marco Guerini .....	3500
<i>WaterJudge: Quality-Detection Trade-off when Watermarking Large Language Models</i> Piotr Molenda, Adian Liusie and Mark Gales .....	3515
<i>Cognitive Overload: Jailbreaking Large Language Models with Overloaded Logical Thinking</i> Nan Xu, Fei Wang, Ben Zhou, Bangzheng Li, Chaowei Xiao and Muhao Chen .....	3526
<i>PAELLA: Parameter-Efficient Lightweight Language-Agnostic Captioning Model</i> Rita Ramos, Emanuele Bugliarello, Bruno Martins and Desmond Elliott .....	3549
<i>OSCaR: Object State Captioning and State Change Representation</i> Nguyen Nguyen, Jing Bi, Ali Vosoughi, Yapeng Tian, Pooyan Fazli and Chenliang Xu .....	3565
<i>SumCSE: Summary as a transformation for Contrastive Learning</i> Raghuveer Thirukovalluru, Xiaolan Wang, Jun Chen, Shuyang Li, Jie Lei, Rong Jin and Bhuwan Dhingra .....	3577
<i>The Curious Decline of Linguistic Diversity: Training Language Models on Synthetic Text</i> Yanzhu Guo, Guokan Shang, Michalis Vazirgiannis and Chloé Clavel .....	3589
<i>PersonaLLM: Investigating the Ability of Large Language Models to Express Personality Traits</i> Hang Jiang, Xiajie Zhang, Xubo Cao, Cynthia Breazeal, Deb Roy and Jad Kabbara .....	3605



<i>FIRE: A Dataset for Financial Relation Extraction</i>	
Hassan Hamad, Abhinav Kumar Thakur, Nijil Kollari, Sujith Pulikodan and Keith M. Chugg	3628
<i>MusiLingo: Bridging Music and Text with Pre-trained Language Models for Music Captioning and Query Response</i>	
Zihao Deng, Yinghao Ma, Yudong Liu, Rongchen Guo, Ge Zhang, Wenhui Chen, Wenhao Huang and Emmanouil Benetos	3643
<i>Investigating Acceleration of LLaMA Inference by Enabling Intermediate Layer Decoding via Instruction Tuning with 'LITE'</i>	
Neeraj Varshney, Agneet Chatterjee, Mihir Parmar and Chitta Baral	3656
<i>Instruction-following Evaluation through Verbalizer Manipulation</i>	
Shiyang Li, Jun Yan, Hai Wang, Zheng Tang, Xiang Ren, Vijay Srinivasan and Hongxia Jin	3678
<i>WebWISE: Unlocking Web Interface Control for LLMs via Sequential Exploration</i>	
Heyi Tao, Sethuraman T V, Michal Shlapentokh-Rothman, Tanmay Gupta, Heng Ji and Derek Hoiem	3693
<i>CodeLM: Aligning Language Models with Tailored Synthetic Data</i>	
Zifeng Wang, Chun-Liang Li, Vincent Perot, Long Le, Jin Miao, Zizhao Zhang, Chen-Yu Lee and Tomas Pfister	3712
<i>Prompting Few-shot Multi-hop Question Generation via Comprehending Type-aware Semantics</i>	
Zefeng Lin, Weidong Chen, Yan Song and Yongdong Zhang	3730
<i>When Hindsight is Not 20/20: Testing Limits on Reflective Thinking in Large Language Models</i>	
Yanhong Li, Chenghao Yang and Allyson Ettinger	3741
<i>CoDa: Constrained Generation based Data Augmentation for Low-Resource NLP</i>	
Chandra Kiran Reddy Evuru, Sreyan Ghosh, Sonal Kumar, Ramaneswaran S, Utkarsh Tyagi and Dinesh Manocha	3754
<i>Synonym relations affect object detection learned on vision-language data</i>	
Giacomo Nebbia and Adriana Kovashka	3770
<i>CM-TTS: Enhancing Real Time Text-to-Speech Synthesis Efficiency through Weighted Samplers and Consistency Models</i>	
Xiang Li, FanBu FanBu, Ambuj Mehrish, Yingting Li, Jiale Han, Bo Cheng and Soujanya Poria	3777
<i>RobustSentEmbed: Robust Sentence Embeddings Using Adversarial Self-Supervised Contrastive Learning</i>	
Javad Rafiei Asl, Prajwal Panzade, Eduardo Blanco, Daniel Takabi and Zhipeng Cai	3795
<i>Characterizing Human and Zero-Shot GPT-3.5 Object-Similarity Judgments</i>	
D Estes McKnight and Alona Fyshe	3810
<i>Self-Demos: Eliciting Out-of-Demonstration Generalizability in Large Language Models</i>	
Wei He, Shichun Liu, Jun Zhao, Yiwen Ding, Yi Lu, Zhiheng Xi, Tao Gui, Qi Zhang and Xuanjing Huang	3829
<i>Getting Sick After Seeing a Doctor? Diagnosing and Mitigating Knowledge Conflicts in Event Temporal Reasoning</i>	
Tianqing Fang, Zhaowei Wang, Wenxuan Zhou, Hongming Zhang, Yangqiu Song and Muhao Chen	3846

<i>MCECR: A Novel Dataset for Multilingual Cross-Document Event Coreference Resolution</i>	
Amir Pouran Ben Veysseh, Viet Dac Lai, Chien Van Nguyen, Franck Dernoncourt and Thien Huu Nguyen .....	3869
<i>Sentiment Analysis in the Era of Large Language Models: A Reality Check</i>	
Wenxuan Zhang, Yue Deng, Bing Liu, Sinno Jialin Pan and Lidong Bing .....	3881
<i>Tokenizer Choice For LLM Training: Negligible or Crucial?</i>	
Mehdi Ali, Michael Fromm, Klaudia Thellmann, Richard Rutmann, Max Lübbering, Johannes Leveling, Katrin Klug, Jan Ebert, Niclas Doll, Jasper Schulze Buschhoff, Charvi Jain, Alexander Arno Weber, Lena Jurkschat, Hammam Abdelwahab, Chelsea Maria John, Pedro Ortiz Suarez, Malte Ostendorff, Samuel Weinbach, Rafet Sifa, Stefan Kesselheim and Nicolas Flores-Herr .....	3907
<i>Think Before You Speak: Cultivating Communication Skills of Large Language Models via Inner Monologue</i>	
Junkai Zhou, Liang Pang, Huawei Shen and Xueqi Cheng .....	3925
<i>The Impact of Differential Privacy on Group Disparity Mitigation</i>	
Victor Petrén Bach Hansen, Atula Tejaswi Neerkaje, Ramit Sawhney, Lucie Flek and Anders Søggaard .....	3952
<i>Isometric Neural Machine Translation using Phoneme Count Ratio Reward-based Reinforcement Learning</i>	
Shivam Ratnakant Mhaskar, Nirmesh J. Shah, Mohammadi Zaki, Ashishkumar Prabhakar Gudmalwar, Pankaj Wasnik and Rajiv Ratn Shah .....	3966
<i>Read between the lines - Functionality Extraction From READMEs</i>	
Prince Kumar, Srikanth G. Tamilselvam and Dinesh Garg .....	3977
<i>AbsPyramid: Benchmarking the Abstraction Ability of Language Models with a Unified Entailment Graph</i>	
Zhaowei Wang, Haochen Shi, Weiqi Wang, Tianqing Fang, Hongming Zhang, Sehyun Choi, Xin Liu and Yangqiu Song .....	3991
<i>Few-TK: A Dataset for Few-shot Scientific Typed Keyphrase Recognition</i>	
Avishek Lahiri, Pratyay Sarkar, Medha Sen, Debarshi Kumar Sanyal and Imon Mukherjee .	4011
<i>Language Models can be Deductive Solvers</i>	
Jiazhan Feng, Ruochen Xu, Junheng Hao, Hiteshi Sharma, yelong shen, Dongyan Zhao and Weizhu Chen .....	4026
<i>Interpreting User Requests in the Context of Natural Language Standing Instructions</i>	
Nikita Moghe, Patrick Xia, Jacob Andreas, Jason Eisner, Benjamin Van Durme and Harsh Jhamtani .....	4043
<i>Secure Your Model: An Effective Key Prompt Protection Mechanism for Large Language Models</i>	
Ruixiang Tang, Yu-Neng Chuang, Xuanting Cai, Mengnan Du and Xia Hu .....	4061
<i>Enhancing Chain-of-Thoughts Prompting with Iterative Bootstrapping in Large Language Models</i>	
Jiashuo Sun, Yi Luo, Yeyun Gong, Chen Lin, yelong shen, Jian Guo and Nan Duan .....	4074
<i>Do Prompt Positions Really Matter?</i>	
Junyu Mao, Stuart E. Middleton and Mahesan Niranjan .....	4102
<i>Natural Language Embedded Programs for Hybrid Language Symbolic Reasoning</i>	
Tianhua Zhang, Jiaxin Ge, Hongyin Luo, Yung-Sung Chuang, Mingye Gao, Yuan Gong, Yoon Kim, Xixin Wu, Helen M. Meng and James R. Glass .....	4131

<i>A Study on Scaling Up Multilingual News Framing Analysis</i> Syeda Sabrina Akter and Antonios Anastasopoulos .....	4156
<i>ViGLUE: A Vietnamese General Language Understanding Benchmark and Analysis of Vietnamese Language Models</i> Minh-Nam Tran, Phu-Vinh Nguyen, Long HB Nguyen and Dien Dinh .....	4174
<i>Exploring the Trade-off Between Model Performance and Explanation Plausibility of Text Classifiers Using Human Rationales</i> Lucas Emanuel Resck, Marcos M. Raimundo and Jorge Poco .....	4190
<i>Unlocking Parameter-Efficient Fine-Tuning for Low-Resource Language Translation</i> Tong Su, Xin Peng, Sarubi Thillainathan, David Guzmán, Surangika Ranathunga and En-Shiun Annie Lee .....	4217
<i>ADaPT: As-Needed Decomposition and Planning with Language Models</i> Archiki Prasad, Alexander Koller, Mareike Hartmann, Peter Clark, Ashish Sabharwal, Mohit Bansal and Tushar Khot .....	4226
<i>Guiding Large Language Models to Post-Edit Machine Translation with Error Annotations</i> Dayeon Ki and Marine Carpuat .....	4253
<i>Non-contrastive sentence representations via self-supervision</i> Duccio Pappadopulo and Marco Farina .....	4274
<i>Semantically-Prompted Language Models Improve Visual Descriptions</i> Michael Ogezi, Bradley Hauer and Grzegorz Kondrak .....	4285
<i>GenTKG: Generative Forecasting on Temporal Knowledge Graph with Large Language Models</i> Ruotong Liao, Xu Jia, Yangzhe Li, Yunpu Ma and Volker Tresp .....	4303
<i>A Transformer with Stack Attention</i> Jiaoda Li, Jennifer C. White, Mrinmaya Sachan and Ryan Cotterell .....	4318
<i>InstructEval: Systematic Evaluation of Instruction Selection Methods</i> Anirudh Ajith, Chris Pan, Mengzhou Xia, Ameet Deshpande and Karthik R Narasimhan ...	4336
<i>RecMind: Large Language Model Powered Agent For Recommendation</i> Yancheng Wang, Ziyang Jiang, Zheng Chen, Fan Yang, Yingxue Zhou, Eunah Cho, Xing Fan, Yanbin Lu, Xiaojiang Huang and Yingzhen Yang .....	4351
<i>GOLD: Generalized Knowledge Distillation via Out-of-Distribution-Guided Language Data Generation</i> Mohsen Gholami, Mohammad Akbari, Tianxi Hu, Vaden Masrani, Z. Jane Wang and Yong Zhang	4365
<i>How Lexical is Bilingual Lexicon Induction?</i> Harsh Kohli, Helian Feng, Nicholas Andrew Dronen, Calvin McCarter, Sina Moeini and Ali Kebari-hotbi .....	4381
<i>Fumbling in Babel: An Investigation into ChatGPT's Language Identification Ability</i> Wei-Rui Chen, Ife Adebara, Khai Duy Doan, Qisheng Liao and Muhammad Abdul-Mageed	4387
<i>Targeted Augmentation for Low-Resource Event Extraction</i> Sijia Wang and Lifu Huang .....	4414
<i>Asking More Informative Questions for Grounded Retrieval</i> Sedrick Keh, Justin T Chiu and Daniel Fried .....	4429

<i>Efficient Citer: Tuning Large Language Models for Enhanced Answer Quality and Verification</i> Marzieh S. Tahaei, Aref Jafari, Ahmad Rashid, David Alfonso-Hermelo, Khalil Bibi, Yimeng Wu, Ali Ghodsi, Boxing Chen and Mehdi Rezagholizadeh . . . . .	4443
<i>Addressing Healthcare-related Racial and LGBTQ+ Biases in Pretrained Language Models</i> Sean Xie, Saeed Hassanpour and Soroush Vosoughi . . . . .	4451
<i>ATG: Benchmarking Automated Theorem Generation for Generative Language Models</i> Xiaohan Lin, Qingxing Cao, Yinya Huang, Zhicheng Yang, Zhengying Liu, Zhenguo Li and Xiaodan Liang . . . . .	4465
<i>Benchmarking Generation and Evaluation Capabilities of Large Language Models for Instruction Controllable Summarization</i> Yixin Liu, Alexander Fabbri, Jiawen Chen, Yilun Zhao, Simeng Han, Shafiq Joty, Pengfei Liu, Dragomir Radev, Chien-Sheng Wu and Arman Cohan . . . . .	4481
<i>NeuroComparatives: Neuro-Symbolic Distillation of Comparative Knowledge</i> Phillip Howard, Junlin Wang, Vasudev Lal, Gadi Singer, Yejin Choi and Swabha Swayamdipta . . . . .	4502
<i>Emotion-Anchored Contrastive Learning Framework for Emotion Recognition in Conversation</i> Fangxu Yu, Junjie Guo, Zhen Wu and Xinyu Dai . . . . .	4521
<i>SUQL: Conversational Search over Structured and Unstructured Data with Large Language Models</i> Shicheng Liu, Jialiang Xu, Wesley Tjangnaka, Sina Semnani, Chen Jie Yu and Monica Lam . . . . .	4535
<i>On Evaluating the Integration of Reasoning and Action in LLM Agents with Database Question Answering</i> Linyong Nan, Ellen Zhang, Weijin Zou, Yilun Zhao, Wenfei Zhou and Arman Cohan . . . . .	4556
<i>CARE: Extracting Experimental Findings From Clinical Literature</i> Aakanksha Naik, Bailey Kuehl, Erin Bransom, Doug Downey and Tom Hope . . . . .	4580
<i>Personalized Federated Learning for Text Classification with Gradient-Free Prompt Tuning</i> Rui Wang, Tong Yu, Ruiyi Zhang, Sungchul Kim, Ryan A. Rossi, Handong Zhao, Junda Wu, Subrata Mitra, Lina Yao and Ricardo Henao . . . . .	4597
<i>SGSH: Stimulate Large Language Models with Skeleton Heuristics for Knowledge Base Question Generation</i> Shasha Guo, Lizi Liao, Jing Zhang, Yanling Wang, Cuiping Li and Hong Chen . . . . .	4613
<i>Biomedical Entity Representation with Graph-Augmented Multi-Objective Transformer</i> Andrey Sakhovskiy, Natalia Semenova, Artur Kadurin and Elena Tutubalina . . . . .	4626
<i>Cross-Lingual Summarization with Pseudo-Label Regularization</i> Thang Le . . . . .	4644
<i>On the Way to Gentle AI Counselor: Politeness Cause Elicitation and Intensity Tagging in Code-mixed Hinglish Conversations for Social Good</i> Priyanshu Priya, gopendra Vikram singh, Mauajama Firdaus, Jyotsna Agrawal and Asif Ekbal . . . . .	4678
<i>Leveraging Summarization for Unsupervised Dialogue Topic Segmentation</i> Alekssei Artemiev, Daniil Parinov, Alexey Grishanov, Ivan Borisov, Alexey Vasilev, Daniil Muravetskii, Aleksey Rezvykh, Aleksei Goncharov and Andrey Savchenko . . . . .	4697

<i>LLaMA-Rider: Spurring Large Language Models to Explore the Open World</i> Yicheng Feng, Yuxuan Wang, jiazheng liu, Sipeng Zheng and Zongqing Lu.....	4705
<i>Contrastive Learning as a Polarizer: Mitigating Gender Bias by Fair and Biased sentences</i> Kyungmin Park, Sihyun Oh, Daehyun Kim and Juae Kim.....	4725
<i>PoLLMgraph: Unraveling Hallucinations in Large Language Models via State Transition Dynamics</i> Derui Zhu, Dingfan Chen, Qing Li, Zongxiong Chen, Lei Ma, Jens Grossklags and Mario Fritz 4737	
<i>Improving Health Question Answering with Reliable and Time-Aware Evidence Retrieval</i> Juraj Vladika and Florian Matthes.....	4752
<i>DecoderLens: Layerwise Interpretation of Encoder-Decoder Transformers</i> Anna Langedijk, Hosein Mohebbi, Gabriele Sarti, Willem Zuidema and Jaap Jumelet.....	4764