

Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)

Singapore
6-10 December 2023

Volume 1 of 21

ISBN: 978-1-7138-8589-4

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

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

Printed with permission by Curran Associates, Inc. (2025)

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>IAG: Induction-Augmented Generation Framework for Answering Reasoning Questions</i> Zhebin Zhang, Xinyu Zhang, Yuanhang Ren, Saijiang Shi, Meng Han, Yongkang Wu, Ruofei Lai and Zhao Cao	1
<i>Absolute Position Embedding Learns Sinusoid-like Waves for Attention Based on Relative Position</i> Yuji Yamamoto and Takuya Matsuzaki	15
<i>Chinese Lexical Substitution: Dataset and Method</i> Jipeng Qiang, Kang Liu, Ying Li, Yun Li, Yi Zhu, Yun-Hao Yuan, Xiaocheng Hu and Xiaoye Ouyang	29
<i>Decoding the Silent Majority: Inducing Belief Augmented Social Graph with Large Language Model for Response Forecasting</i> Chenkai Sun, Jinning Li, Yi Fung, Hou Pong Chan, Tarek Abdelzaher, ChengXiang Zhai and Heng Ji	43
<i>Fine-grained Conversational Decoding via Isotropic and Proximal Search</i> Yuxuan Yao, Han Wu, Qiling Xu and Linqi Song	58
<i>Holistic Inter-Annotator Agreement and Corpus Coherence Estimation in a Large-scale Multilingual Annotation Campaign</i> Nicolas Stefanovitch and Jakub Piskorski	71
<i>PHD: Pixel-Based Language Modeling of Historical Documents</i> Nadav Borenstein, Phillip Rust, Desmond Elliott and Isabelle Augenstein	87
<i>Primacy Effect of ChatGPT</i> Yiwei Wang, Yujun Cai, Muhao Chen, Yuxuan Liang and Bryan Hooi	108
<i>Evaluating the Rationale Understanding of Critical Reasoning in Logical Reading Comprehension</i> Akira Kawabata and Saku Sugawara	116
<i>Evaluating and Modeling Attribution for Cross-Lingual Question Answering</i> Benjamin Muller, John Frederick Wieting, Jonathan H. Clark, Tom Kwiatkowski, Sebastian Ru- der, Livio Baldini Soares, Roei Aharoni, Jonathan Herzig and Xinyi Wang	144
<i>Better Quality Pre-training Data and T5 Models for African Languages</i> Akintunde Oladipo, Mofetoluwa Adeyemi, Orevaoghene Ahia, Abraham Toluwase Owodunni, Odunayo Ogundepo, David Ifeoluwa Adelani and Jimmy Lin	158
<i>Sparse Universal Transformer</i> Shawn Tan, Yikang Shen, Zhenfang Chen, Aaron Courville and Chuang Gan	169
<i>Theory of Mind for Multi-Agent Collaboration via Large Language Models</i> Huaoli, Yu Quan Chong, Simon Stepputtis, Joseph Campbell, Dana Hughes, Charles Michael Lewis and Katia P. Sycara	180
<i>Establishing Trustworthiness: Rethinking Tasks and Model Evaluation</i> Robert Litschko, Max Müller-Eberstein, Rob van der Goot, Leon Weber-Genzel and Barbara Plank 193	

<i>Let's Think Frame by Frame with VIP: A Video Infilling and Prediction Dataset for Evaluating Video Chain-of-Thought</i>	
Vaishnavi Himakunthala, Andy Ouyang, Daniel Philip Rose, Ryan He, Alex Mei, Yujie Lu, Chinmay Sonar, Michael Saxon and William Yang Wang	204
<i>GPTAraEval: A Comprehensive Evaluation of ChatGPT on Arabic NLP</i>	
Md Tawkat Islam Khondaker, Abdul Waheed, El Moatez Billah Nagoudi and Muhammad Abdul-Mageed	220
<i>Dual-Channel Span for Aspect Sentiment Triplet Extraction</i>	
Pan Li, Ping Li and Kai Zhang	248
<i>Cultural Concept Adaptation on Multimodal Reasoning</i>	
Zhi Li and Yin Zhang	262
<i>Understanding Compositional Data Augmentation in Typologically Diverse Morphological Inflection</i>	
Farhan Samir and Miikka Silfverberg	277
<i>Evaluating Object Hallucination in Large Vision-Language Models</i>	
Yifan Li, Yifan Du, Kun Zhou, Jinpeng Wang, Xin Zhao and Ji-Rong Wen	292
<i>Event Ontology Completion with Hierarchical Structure Evolution Networks</i>	
Pengfei Cao, Yupu Hao, Yubo Chen, Kang Liu, Jiexin Xu, Huaijun Li, Xiaojian Jiang and Jun Zhao	306
<i>Parameter-efficient Tuning for Large Language Model without Calculating Its Gradients</i>	
Feihu Jin, Jiajun Zhang and Chengqing Zong	321
<i>Discourse Structures Guided Fine-grained Propaganda Identification</i>	
Yuanyuan Lei and Ruihong Huang	331
<i>CompoundPiece: Evaluating and Improving Decomposing Performance of Language Models</i>	
Benjamin Minixhofer, Jonas Pfeiffer and Ivan Vulić	343
<i>Improving Image Captioning via Predicting Structured Concepts</i>	
Ting Wang, Weidong Chen, Yuanhe Tian, Yan Song and Zhendong Mao	360
<i>GATITOS: Using a New Multilingual Lexicon for Low-resource Machine Translation</i>	
Alexander Jones, Isaac Rayburn Caswell, Orhan Firat and Ishank Saxena	371
<i>Continually Improving Extractive QA via Human Feedback</i>	
Ge Gao, Hung-Ting Chen, Yoav Artzi and Eunsol Choi	406
<i>Using Interpretation Methods for Model Enhancement</i>	
Zhuo Chen, Chengyue Jiang and Kewei Tu	424
<i>An Expression Tree Decoding Strategy for Mathematical Equation Generation</i>	
Wenqi Zhang, Yongliang Shen, Qingpeng Nong, Zeqi Tan, Yanna Ma and Weiming Lu	439
<i>Bootstrapping Small & High Performance Language Models with Unmasking-Removal Training Policy</i>	
Yahan Yang, Elior Sulem, Insup Lee and Dan Roth	457
<i>Diversity Enhanced Narrative Question Generation for Storybooks</i>	
Hokeun Yoon and JinYeong Bak	465
<i>Debiasing Made State-of-the-art: Revisiting the Simple Seed-based Weak Supervision for Text Classification</i>	
Chengyu Dong, Zihan Wang and Jingbo Shang	483

<i>How to Enhance Causal Discrimination of Utterances: A Case on Affective Reasoning</i>	
Hang Chen, Xinyu Yang, Jing Luo and Wenjing Zhu	494
<i>Compressing and Debiasing Vision-Language Pre-Trained Models for Visual Question Answering</i>	
Qingyi Si, Yuanxin Liu, Zheng Lin, Peng Fu, Yanan Cao and Weiping Wang	513
<i>Selectively Answering Ambiguous Questions</i>	
Jeremy R. Cole, Michael JQ Zhang, Daniel Gillick, Julian Martin Eisenschlos, Bhuwan Dhingra and Jacob Eisenstein	530
<i>Temporal Knowledge Graph Forecasting Without Knowledge Using In-Context Learning</i>	
Dong-Ho Lee, Kian Ahrabian, Woojeong Jin, Fred Morstatter and Jay Pujara	544
<i>Knowledge Graph Compression Enhances Diverse Commonsense Generation</i>	
EunJeong Hwang, Veronika Thost, Vered Shwartz and Tengfei Ma	558
<i>Pragmatic Reasoning Unlocks Quantifier Semantics for Foundation Models</i>	
Yiyuan Li, Rakesh R Menon, Sayan Ghosh and Shashank Srivastava	573
<i>LLM-FP4: 4-Bit Floating-Point Quantized Transformers</i>	
Shih-yang Liu, Zechun Liu, Xijie Huang, Pingcheng Dong and Kwang-Ting Cheng	592
<i>Improving Biomedical Abstractive Summarisation with Knowledge Aggregation from Citation Papers</i>	
Chen Tang, Shun Wang, Tomas Goldsack and Chenghua Lin	606
<i>Explanation Selection Using Unlabeled Data for Chain-of-Thought Prompting</i>	
Xi Ye and Greg Durrett	619
<i>HalOmi: A Manually Annotated Benchmark for Multilingual Hallucination and Omission Detection in Machine Translation</i>	
David Dale, Elena Voita, Janice Lam, Prangthip Hansanti, Christophe Ropers, Elahe Kalbassi, Cynthia Gao, Loic Barrault and Marta R. Costa-jussà	638
<i>Gradient-based Gradual Pruning for Language-Specific Multilingual Neural Machine Translation</i>	
Dan He, Minh-Quang Pham, Thanh-Le Ha and Marco Turchi	654
<i>LLM-powered Data Augmentation for Enhanced Cross-lingual Performance</i>	
Chenxi Whitehouse, Monojit Choudhury and Alham Fikri Aji	671
<i>Prompt-based Logical Semantics Enhancement for Implicit Discourse Relation Recognition</i>	
Chenxu Wang, Ping Jian and Mu Huang	687
<i>VLIS: Unimodal Language Models Guide Multimodal Language Generation</i>	
Jiwan Chung and Youngjae Yu	700
<i>Conceptual structure coheres in human cognition but not in large language models</i>	
Siddharth Suresh, Kushin Mukherjee, Xizheng Yu, Wei-Chun Huang, Lisa Padua and Timothy T. Rogers	722
<i>Towards LLM-driven Dialogue State Tracking</i>	
Yujie Feng, Zexin Lu, Bo Liu, Liming Zhan and Xiao-Ming Wu	739
<i>Learning Language-guided Adaptive Hyper-modality Representation for Multimodal Sentiment Analysis</i>	
Haoyu Zhang, Yu Wang, Guanghao Yin, Kejun Liu, Yuanyuan Liu and Tianshu Yu	756

<i>Multitask Multimodal Prompted Training for Interactive Embodied Task Completion</i>	
Georgios Pantazopoulos, Malvina Nikandrou, Amit Parekh, Bhathiya Hemanthage, Arash Eshghi, Ioannis Konstas, Verena Rieser, Oliver Lemon and Alessandro Suglia	768
<i>We're Afraid Language Models Aren't Modeling Ambiguity</i>	
Alisa Liu, Zhaofeng Wu, Julian Michael, Alane Suhr, Peter West, Alexander Koller, Swabha Swayamdipta, Noah A. Smith and Yejin Choi	790
<i>Linear-Time Modeling of Linguistic Structure: An Order-Theoretic Perspective</i>	
Tianyu Liu, Afra Amini, Mrinmaya Sachan and Ryan Cotterell	808
<i>GEMINI: Controlling The Sentence-Level Summary Style in Abstractive Text Summarization</i>	
Guangsheng Bao, Zebin Ou and Yue Zhang	831
<i>Fidelity-Enriched Contrastive Search: Reconciling the Faithfulness-Diversity Trade-Off in Text Generation</i>	
Wei-Lin Chen, Cheng-Kuang Wu, Hsin-Hsi Chen and Chung-Chi Chen	843
<i>Analyzing Norm Violations in Live-Stream Chat</i>	
Jihyung Moon, Dong-Ho Lee, Hyundong Justin Cho, Woojeong Jin, Chan Young Park, Minwoo Kim, Jonathan May, Jay Pujara and Sungjoon Park	852
<i>Coarse-to-Fine Contrastive Learning in Image-Text-Graph Space for Improved Vision-Language Compositionality</i>	
Harman Singh, Pengchuan Zhang, Qifan Wang, Mengjiao Wang, Wenhan Xiong, Jingfei Du and Yu Chen	869
<i>Reading Books is Great, But Not if You Are Driving! Visually Grounded Reasoning about Defeasible Commonsense Norms</i>	
Seungju Han, Junhyeok Kim, Jack Hessel, Liwei Jiang, Jiwan Chung, Yejin Son, Yejin Choi and Youngjae Yu	894
<i>Enhancing Uncertainty-Based Hallucination Detection with Stronger Focus</i>	
Tianhang Zhang, Lin Qiu, Qipeng Guo, Cheng Deng, Yue Zhang, Zheng Zhang, Chenghu Zhou, Xinbing Wang and Luoyi Fu	915
<i>FactKB: Generalizable Factuality Evaluation using Language Models Enhanced with Factual Knowledge</i>	
Shangbin Feng, Vidhisha Balachandran, Yuyang Bai and Yulia Tsvetkov	933
<i>Mitigating Backdoor Poisoning Attacks through the Lens of Spurious Correlation</i>	
Xuanli He, Qionghai Xu, Jun Wang, Benjamin I. P. Rubinstein and Trevor Cohn	953
<i>Symbol tuning improves in-context learning in language models</i>	
Jerry Wei, Le Hou, Andrew Kyle Lampinen, Xiangning Chen, Da Huang, Yi Tay, Xinyun Chen, Yifeng Lu, Denny Zhou, Tengyu Ma and Quoc V Le	968
<i>The neural dynamics of word recognition and integration</i>	
Jon Gauthier and Roger P. Levy	980
<i>Tree of Clarifications: Answering Ambiguous Questions with Retrieval-Augmented Large Language Models</i>	
Gangwoo Kim, Sungdong Kim, Byeongguk Jeon, Joonsuk Park and Jaewoo Kang	996
<i>Incorporating Worker Perspectives into MTurk Annotation Practices for NLP</i>	
Olivia Huang, Eve Fleisig and Dan Klein	1010

<i>Predict the Future from the Past? On the Temporal Data Distribution Shift in Financial Sentiment Classifications</i>	
Yue Guo, Chenxi Hu and Yi Yang	1029
<i>Look-back Decoding for Open-Ended Text Generation</i>	
Nan Xu, Chunting Zhou, Asli Celikyilmaz and Xuezhe Ma	1039
<i>Large Language Models Can Self-Improve</i>	
Jiaxin Huang, Shixiang Shane Gu, Le Hou, Yuexin Wu, Xuezhi Wang, Hongkun Yu and Jiawei Han	1051
<i>CodeT5+: Open Code Large Language Models for Code Understanding and Generation</i>	
Yue Wang, Hung Le, Akhilesh Deepak Gotmare, Nghi D. Q. Bui, Junnan Li and Steven Hoi	1069
<i>Structural generalization in COGS: Supertagging is (almost) all you need</i>	
Alban Petit, Caio Filippo Corro and François Yvon	1089
<i>BioT5: Enriching Cross-modal Integration in Biology with Chemical Knowledge and Natural Language Associations</i>	
Qizhi Pei, Wei Zhang, Jinhua Zhu, Kehan Wu, Kaiyuan Gao, Lijun Wu, Yingce Xia and Rui Yan	1102
<i>Hyperpolyglot LLMs: Cross-Lingual Interpretability in Token Embeddings</i>	
Andrea W Wen-Yi and David Mimno	1124
<i>Target-oriented Proactive Dialogue Systems with Personalization: Problem Formulation and Dataset Curation</i>	
Jian Wang, Yi Cheng, Dongding Lin, Chak Tou Leong and Wenjie Li	1132
<i>SeqXGPT: Sentence-Level AI-Generated Text Detection</i>	
Pengyu Wang, Linyang Li, Ke Ren, Botian Jiang, Dong Zhang and Xipeng Qiu	1144
<i>QTSumm: Query-Focused Summarization over Tabular Data</i>	
Yilun Zhao, Zhenting Qi, Linyong Nan, Boyu Mi, Yixin Liu, Weijin Zou, Simeng Han, Ruizhe Chen, Xiangru Tang, Yumo Xu, Dragomir Radev and Arman Cohan	1157
<i>From Wrong To Right: A Recursive Approach Towards Vision-Language Explanation</i>	
Jiaxin Ge, Sanjay Subramanian, Trevor Darrell and Boyi Li	1173
<i>‘Don’t Get Too Technical with Me’: A Discourse Structure-Based Framework for Automatic Science Journalism</i>	
Ronald Cardenas, Bingsheng Yao, Dakuo Wang and Yufang Hou	1186
<i>LACMA: Language-Aligning Contrastive Learning with Meta-Actions for Embodied Instruction Following</i>	
Cheng-Fu Yang, Yen-Chun Chen, Jianwei Yang, Xiyang Dai, Lu Yuan, Yu-Chiang Frank Wang and Kai-Wei Chang	1203
<i>Penalty Decoding: Well Suppress the Self-Reinforcement Effect in Open-Ended Text Generation</i>	
Wenhong Zhu, Hongkun Hao and Rui Wang	1218
<i>Towards Robust Pruning: An Adaptive Knowledge-Retention Pruning Strategy for Language Models</i>	
Jianwei Li, Qi Lei, Wei Cheng and Dongkuan Xu	1229
<i>Clinical Contradiction Detection</i>	
Dave Makhervaks, Plia Gillis and Kira Radinsky	1248

<i>Vera: A General-Purpose Plausibility Estimation Model for Commonsense Statements</i> Jiacheng Liu, Wenya Wang, Dianzhuo Wang, Noah A. Smith, Yejin Choi and Hannaneh Hajishirzi	1264
<i>Text-Transport: Toward Learning Causal Effects of Natural Language</i> Victoria Lin, Louis-Philippe Morency and Eli Ben-Michael	1288
<i>How Does Generative Retrieval Scale to Millions of Passages?</i> Ronak Pradeep, Kai Hui, Jai Gupta, Adam D Lelkes, Honglei Zhuang, Jimmy Lin, Donald Metzler and Vinh Q. Tran	1305
<i>Unveiling the Implicit Toxicity in Large Language Models</i> Jiaxin Wen, Pei Ke, Hao Sun, Zhixin Zhang, Chengfei Li, Jinfeng Bai and Minlie Huang	1322
<i>Is ChatGPT a General-Purpose Natural Language Processing Task Solver?</i> Chengwei Qin, Aston Zhang, Zhuosheng Zhang, Jiaao Chen, Michihiro Yasunaga and Diyi Yang	1339
<i>Length is a Curse and a Blessing for Document-level Semantics</i> Chenghao Xiao, Yizhi Li, G Thomas Hudson, Chenghua Lin and Noura Al Moubayed	1385
<i>ALCUNA: Large Language Models Meet New Knowledge</i> Xunjian Yin, Baizhou Huang and Xiaojun Wan	1397
<i>Location-Aware Visual Question Generation with Lightweight Models</i> Nicholas Collin Suwono, Justin Chen, Tun Min Hung, Ting-Hao Kenneth Huang, I-Bin Liao, Yung-Hui Li, Lun-Wei Ku and Shao-Hua Sun	1415
<i>MemeCap: A Dataset for Captioning and Interpreting Memes</i> EunJeong Hwang and Vered Shwartz	1433
<i>Where to start? Analyzing the potential value of intermediate models</i> Leshem Choshen, Elad Venezian, Shachar Don-Yehiya, Noam Slonim and Yoav Katz	1446
<i>Transcending Scaling Laws with 0.1% Extra Compute</i> Yi Tay, Jason Wei, Hyung Won Chung, Vinh Q. Tran, David So, Siamak Shakeri, Xavier Garcia, Steven Zheng, Jinfeng Rao, Aakanksha Chowdhery, Denny Zhou, Donald Metzler, Slav Petrov, Neil Houlsby, Quoc V Le and Mostafa Dehghani	1471
<i>CoAnnotating: Uncertainty-Guided Work Allocation between Human and Large Language Models for Data Annotation</i> Minzhi Li, Taiwei Shi, Caleb Ziems, Min-Yen Kan, Nancy F. Chen, Zhengyuan Liu and Diyi Yang	1487
<i>Optimizing Retrieval-augmented Reader Models via Token Elimination</i> Moshe Berchansky, Peter Izsak, Avi Caciularu, Ido Dagan and Moshe Wasserblat	1506
<i>WSDMS: Debunk Fake News via Weakly Supervised Detection of Misinforming Sentences with Contextualized Social Wisdom</i> Ruichao Yang, Wei Gao, Jing Ma, Hongzhan Lin and Zhiwei Yang	1525
<i>Robust Prompt Optimization for Large Language Models Against Distribution Shifts</i> Moxin Li, Wenjie Wang, Fuli Feng, Yixin Cao, Jizhi Zhang and Tat-Seng Chua	1539
<i>Exploiting Asymmetry for Synthetic Training Data Generation: SynthIE and the Case of Information Extraction</i> Martin Josifoski, Marija Sakota, Maxime Peyrard and Robert West	1555

<i>Condensing Multilingual Knowledge with Lightweight Language-Specific Modules</i>	
Haoran Xu, Weiting Tan, Shuyue Stella Li, Yunmo Chen, Benjamin Van Durme, Philipp Koehn and Kenton Murray	1575
<i>The Framework Tax: Disparities Between Inference Efficiency in NLP Research and Deployment</i>	
Jared Fernandez, Jacob Kahn, Clara Na, Yonatan Bisk and Emma Strubell	1588
<i>Evaluating Cross-Domain Text-to-SQL Models and Benchmarks</i>	
Mohammadreza Pourreza and Davood Rafiei	1601
<i>Increasing Coverage and Precision of Textual Information in Multilingual Knowledge Graphs</i>	
Simone Conia, Min Li, Daniel Lee, Umar Farooq Minhas, Ihab Ilyas and Yunyao Li	1612
<i>Memory-Based Invariance Learning for Out-of-Domain Text Classification</i>	
Chen Jia and Yue Zhang	1635
<i>Outlier Suppression+: Accurate quantization of large language models by equivalent and effective shifting and scaling</i>	
Xiuying Wei, Yunchen Zhang, Yuhang Li, Xiangguo Zhang, Ruihao Gong, Jinyang Guo and Xianglong Liu	1648
<i>Three Stream Based Multi-level Event Contrastive Learning for Text-Video Event Extraction</i>	
Jiaqi Li, Chuanyi Zhang, Miaozen Du, Dehai Min, Yongrui Chen and Guilin Qi	1666
<i>Diversify Question Generation with Retrieval-Augmented Style Transfer</i>	
Qi Gou, Zehua Xia, Bowen Yu, Haiyang Yu, Fei Huang, Yongbin Li and Nguyen Cam-Tu ..	1677
<i>Fast and Accurate Factual Inconsistency Detection Over Long Documents</i>	
Barrett Martin Lattimer, Patrick CHen, Xinyuan Zhang and Yi Yang	1691
<i>Interpreting Embedding Spaces by Conceptualization</i>	
Adi Simhi and Shaul Markovitch	1704
<i>Knowledge-Augmented Language Model Verification</i>	
Jinheon Baek, Soyeong Jeong, Minki Kang, Jong C. Park and Sung Ju Hwang	1720
<i>A Generation-based Deductive Method for Math Word Problems</i>	
Yuxuan Hu, Jing Zhang, Haoyang Li, Cuiping Li and Hong Chen	1737
<i>Failures Pave the Way: Enhancing Large Language Models through Tuning-free Rule Accumulation</i>	
Zeyuan Yang, Peng Li and Yang Liu	1751
<i>Building Persona Consistent Dialogue Agents with Offline Reinforcement Learning</i>	
Ryan Shea and Zhou Yu	1778
<i>Augmenting Zero-Shot Dense Retrievers with Plug-in Mixture-of-Memories</i>	
Suyu Ge, Chenyan Xiong, Corby Rosset, Arnold Overwijk, Jiawei Han and Paul N. Bennett	1796
<i>Active Instruction Tuning: Improving Cross-Task Generalization by Training on Prompt Sensitive Tasks</i>	
Po-Nien Kung, Fan Yin, Di Wu, Kai-Wei Chang and Nanyun Peng	1813
<i>Towards Example-Based NMT with Multi-Levenshtein Transformers</i>	
Maxime Bouthors, Josep Crego and François Yvon	1830
<i>DUnE: Dataset for Unified Editing</i>	
Afra Feyza Akyürek, Eric L Pan, Garry Kuwanto and Derry Wijaya	1847

<i>“Fifty Shades of Bias”: Normative Ratings of Gender Bias in GPT Generated English Text</i>	
Rishav Hada, Agrima Seth, Harshita Diddee and Kalika Bali	1862
<i>Hybrid Inverted Index Is a Robust Accelerator for Dense Retrieval</i>	
Peitian Zhang, Zheng Liu, Shitao Xiao, Zhicheng Dou and Jing Yao	1877
<i>ChatGPT to Replace Crowdsourcing of Paraphrases for Intent Classification: Higher Diversity and Comparable Model Robustness</i>	
Jan Cegin, Jakub Simko and Peter Brusilovsky	1889
<i>Query-as-context Pre-training for Dense Passage Retrieval</i>	
Xing W, Guangyuan Ma, Wanhui Qian, Zijia Lin and Songlin Hu	1906
<i>A Suite of Generative Tasks for Multi-Level Multimodal Webpage Understanding</i>	
Andrea Burns, Krishna Srinivasan, Joshua Ainslie, Geoff Brown, Bryan A. Plummer, Kate Saenko, Jianmo Ni and Mandy Guo	1917
<i>Democratizing Reasoning Ability: Tailored Learning from Large Language Model</i>	
Zhaoyang Wang, Shaohan Huang, Yuxuan Liu, Jiahai Wang, Minghui Song, Zihan Zhang, Haizhen Huang, Furu Wei, Weiwei Deng, Feng Sun and Qi Zhang	1948
<i>OpenAsp: A Benchmark for Multi-document Open Aspect-based Summarization</i>	
Shmuel Amar, Liat Schiff, Ori Ernst, Asi Shefer, Ori Shapira and Ido Dagan	1967
<i>PEFTDebias : Capturing debiasing information using PEFTs</i>	
Sumit Agarwal, Aditya Srikanth Veerubhotla and Srijan Bansal	1992
<i>Byte Pair Encoding for Symbolic Music</i>	
Nathan Fradet, Nicolas Gutowski, Fabien Chhel and Jean-Pierre Briot	2001
<i>Combining Denoising Autoencoders with Contrastive Learning to fine-tune Transformer Models</i>	
Alejo Lopez-Avila and Víctor Suárez-Paniagua	2021
<i>Self-Influence Guided Data Reweighting for Language Model Pre-training</i>	
Megh Thakkar, Tolga Bolukbasi, Sriram Ganapathy, Shikhar Vashishth, Sarath Chandar and Partha Talukdar	2033
<i>ACTOR: Active Learning with Annotator-specific Classification Heads to Embrace Human Label Variation</i>	
Xinpeng Wang and Barbara Plank	2046
<i>TrueTeacher: Learning Factual Consistency Evaluation with Large Language Models</i>	
Zorik Gekhman, Jonathan Herzig, Roei Aharoni, Chen Elkind and Idan Szpektor	2053
<i>VivesDebate-Speech: A Corpus of Spoken Argumentation to Leverage Audio Features for Argument Mining</i>	
Ramon Ruiz-Dolz and Javier Iranzo Sanchez	2071
<i>Tagging-Assisted Generation Model with Encoder and Decoder Supervision for Aspect Sentiment Triplet Extraction</i>	
Luo Xianlong, Meng Yang and Yihao Wang	2078
<i>Larger Probes Tell a Different Story: Extending Psycholinguistic Datasets Via In-Context Learning</i>	
Namrata Shivagunde, Vladislav Lialin and Anna Rumshisky	2094
<i>Norm of Word Embedding Encodes Information Gain</i>	
Momose Oyama, Sho Yokoi and Hidetoshi Shimodaira	2108

<i>CRT-QA: A Dataset of Complex Reasoning Question Answering over Tabular Data</i> Zhehao Zhang, Xitao Li, Yan Gao and Jian-Guang Lou	2131
<i>Promoting Topic Coherence and Inter-Document Consorts in Multi-Document Summarization via Simplicial Complex and Sheaf Graph</i> Yash Kumar Atri, Arun Iyer, Tanmoy Chakraborty and Vikram Goyal	2154
<i>MAGNIFICo: Evaluating the In-Context Learning Ability of Large Language Models to Generalize to Novel Interpretations</i> Arkil Patel, Satwik Bhattamishra, Siva Reddy and Dzmitry Bahdanau	2167
<i>Generating and Evaluating Tests for K-12 Students with Language Model Simulations: A Case Study on Sentence Reading Efficiency</i> Eric Zelikman, Wanjing Anya Ma, Jasmine Elizabeth Tran, Diyi Yang, Jason D Yeatman and Nick Haber	2190
<i>Counter Turing Test (CT2): AI-Generated Text Detection is Not as Easy as You May Think - Introducing AI Detectability Index (ADI)</i> Megha Chakraborty, S.M Towhidul Islam Tonmoy, S M Mehedi Zaman, Shreya Gautam, Tanay Kumar, Krish Sharma, Niyar R Barman, Chandan Gupta, Vinija Jain, Aman Chadha, Amit P. Sheth and Amitava Das	2206
<i>Revisiting the Optimality of Word Lengths</i> Tiago Pimentel, Clara Meister, Ethan Wilcox, Kyle Mahowald and Ryan Cotterell	2240
<i>Document-level Relationship Extraction by Bidirectional Constraints of Beta Rules</i> Yichun Liu, Zizhong Zhu, Xiaowang Zhang, Zhiyong Feng, Daoqi Chen and Yaxin Li	2256
<i>Instructed Language Models with Retrievers Are Powerful Entity Linkers</i> Zilin Xiao, Ming Gong, Jie Wu, Xingyao Zhang, Linjun Shou and Daxin Jiang	2267
<i>Towards Noise-Tolerant Speech-Referring Video Object Segmentation: Bridging Speech and Text</i> Xiang Li, Jinglu Wang, Xiaohao Xu, Muqiao Yang, Fan Yang, Yizhou Zhao, Rita Singh and Bhiksha Raj	2283
<i>PROSE: A Pronoun Omission Solution for Chinese-English Spoken Language Translation</i> Ke Wang, Xiutian Zhao, Yanghui Li and Wei Peng	2297
<i>A Diachronic Analysis of Paradigm Shifts in NLP Research: When, How, and Why?</i> Aniket Pramanick, Yufang Hou, Saif M. Mohammad and Iryna Gurevych	2312
<i>Does the Correctness of Factual Knowledge Matter for Factual Knowledge-Enhanced Pre-trained Language Models?</i> Boxi Cao, Qiaoyu Tang, Hongyu Lin, Xianpei Han and Le Sun	2327
<i>Syntactic Substitutability as Unsupervised Dependency Syntax</i> Jasper Jian and Siva Reddy	2341
<i>MProto: Multi-Prototype Network with Denoised Optimal Transport for Distantly Supervised Named Entity Recognition</i> Shuhui Wu, Yongliang Shen, Zeqi Tan, Wenqi Ren, Jietian Guo, Shiliang Pu and Weiming Lu	2361
<i>The Shifted and The Overlooked: A Task-oriented Investigation of User-GPT Interactions</i> Siru Ouyang, Shuohang Wang, Yang Liu, Ming Zhong, Yizhu Jiao, Dan Iter, Reid Pryzant, Chenguang Zhu, Heng Ji and Jiawei Han	2375

<i>Learning the Visualness of Text Using Large Vision-Language Models</i>	
Gaurav Verma, Ryan A. Rossi, Christopher Tensmeyer, Jiuxiang Gu and Ani Nenkova	2394
<i>The Past, Present and Better Future of Feedback Learning in Large Language Models for Subjective Human Preferences and Values</i>	
Hannah Rose Kirk, Andrew Michael Bean, Bertie Vidgen, Paul Rottger and Scott A. Hale . .	2409
<i>TempTabQA: Temporal Question Answering for Semi-Structured Tables</i>	
Vivek Gupta, Pranshu Kandoi, Mahek Bhavesh Vora, Shuo Zhang, Yujie He, Ridho Reinanda and Vivek Srikumar	2431
<i>Task-Level Thinking Steps Help Large Language Models for Challenging Classification Task</i>	
Chunhui Du, Jidong Tian, Haoran Liao, Jindou Chen, Hao He and Yaohui Jin	2454
<i>RepoCoder: Repository-Level Code Completion Through Iterative Retrieval and Generation</i>	
Fengji Zhang, Bei Chen, Yue Zhang, Jacky Keung, Jin Liu, Daoguang Zan, Yi Mao, Jian-Guang Lou and Weizhu Chen	2471
<i>Influence Scores at Scale for Efficient Language Data Sampling</i>	
Nikhil Anand, Joshua Tan and Maria Minakova	2485
<i>G-Eval: NLG Evaluation using Gpt-4 with Better Human Alignment</i>	
Yang Liu, Dan Iter, Yichong Xu, Shuohang Wang, Ruochen Xu and Chenguang Zhu	2511
<i>Learning Retrieval Augmentation for Personalized Dialogue Generation</i>	
Qiushi Huang, Shuai Fu, Xubo Liu, Wenwu Wang, Tom Ko, Yu Zhang and Lilian Tang . . .	2523
<i>The Troubling Emergence of Hallucination in Large Language Models - An Extensive Definition, Quantification, and Prescriptive Remediations</i>	
Vipula Rawte, Swagata Chakraborty, Agnibh Pathak, Anubhav Sarkar, S.M Towhidul Islam Tonmoy, Aman Chadha, Amit P. Sheth and Amitava Das	2541
<i>NAIL: Lexical Retrieval Indices with Efficient Non-Autoregressive Decoders</i>	
Livio Baldini Soares, Daniel Gillick, Jeremy R. Cole and Tom Kwiatkowski	2574
<i>Analyzing Modular Approaches for Visual Question Decomposition</i>	
Apoorv Khandelwal, Ellie Pavlick and Chen Sun	2590
<i>Improving Summarization with Human Edits</i>	
Zonghai Yao, Benjamin J Schloss and Sai P Selvaraj	2604
<i>Did You Mean...? Confidence-based Trade-offs in Semantic Parsing</i>	
Elias Stengel-Eskin and Benjamin Van Durme	2621
<i>The Skipped Beat: A Study of Sociopragmatic Understanding in LLMs for 64 Languages</i>	
Chiyu Zhang, Khai Duy Doan, Qisheng Liao and Muhammad Abdul-Mageed	2630
<i>Understanding the Effect of Model Compression on Social Bias in Large Language Models</i>	
Gustavo Gonalves and Emma Strubell	2663
<i>BioPlanner: Automatic Evaluation of LLMs on Protocol Planning in Biology</i>	
Odhran O'Donoghue, Aleksandar Shtedritski, John Ginger, Ralph Abboud, Ali Essam Ghareeb and Samuel G Rodriques	2676
<i>Cross-lingual Prompting: Improving Zero-shot Chain-of-Thought Reasoning across Languages</i>	
Libo Qin, Qiguang Chen, Fuxuan Wei, Shijue Huang and Wanxiang Che	2695

<i>FinGPT: Large Generative Models for a Small Language</i>	
Risto Luukkonen, Ville Komulainen, Jouni Luoma, Anni Eskelinen, Jenna Kanerva, Hanna-Mari Kristiina Kupari, Filip Ginter, Veronika Laippala, Niklas Muennighoff, Aleksandra Piktus, Thomas Wang, Nouamane Tazi, Teven Le Scao, Thomas Wolf, Osmo Suominen, Samuli Sairanen, Mikko Merioksa, Jyrki Heinonen, Aija Vahtola, Samuel Antao and Sampo Pyysalo	2710
<i>Boosting Summarization with Normalizing Flows and Aggressive Training</i>	
Yu Yang and Xiaotong Shen	2727
<i>Indicative Summarization of Long Discussions</i>	
Shahbaz Syed, Dominik Schwabe, Khalid Al Khatib and Martin Potthast	2752
<i>A Framework for Vision-Language Warm-up Tasks in Multimodal Dialogue Models</i>	
Jaewook Lee, Seongsik Park, Seong-Heum Park, Hongjin Kim and Harksoo Kim	2789
<i>Once is Enough: A Light-Weight Cross-Attention for Fast Sentence Pair Modeling</i>	
Yuanhang Yang, Shiyi Qi, Chuanyi Liu, Qifan Wang, Cuiyun Gao and Zenglin Xu	2800
<i>Plan, Verify and Switch: Integrated Reasoning with Diverse X-of-Thoughts</i>	
Tengxiao Liu, Qipeng Guo, Yuqing Yang, Xiangkun Hu, Yue Zhang, Xipeng Qiu and Zheng Zhang	2807
<i>GLEN: General-Purpose Event Detection for Thousands of Types</i>	
Sha Li, Qiusi Zhan, Kathryn Conger, Martha Palmer, Heng Ji and Jiawei Han	2823
<i>Hierarchical Pretraining on Multimodal Electronic Health Records</i>	
Xiaochen Wang, Junyu Luo, Jiaqi Wang, Ziyi Yin, Suhan Cui, Yuan Zhong, Yaqing Wang and Fenglong Ma	2839
<i>Critic-Driven Decoding for Mitigating Hallucinations in Data-to-text Generation</i>	
Mateusz Lango and Ondrej Dusek	2853
<i>Bridging the Gap between Synthetic and Authentic Images for Multimodal Machine Translation</i>	
Wenyu Guo, Qingkai Fang, Dong Yu and Yang Feng	2863
<i>DEPN: Detecting and Editing Privacy Neurons in Pretrained Language Models</i>	
Xinwei Wu, Junzhuo Li, Minghui Xu, Weilong Dong, Shuangzhi Wu, Chao Bian and Deyi Xiong	2875
<i>Investigating Bias in Multilingual Language Models: Cross-Lingual Transfer of Debiasing Techniques</i>	
Manon Reusens, Philipp Borchert, Margot Mieskes, Jochen De Weerd and Bart Baesens	2887
<i>Can Language Models Laugh at YouTube Short-form Videos?</i>	
Dayoon Ko, Sangho Lee and Gunhee Kim	2897
<i>Random Entity Quantization for Parameter-Efficient Compositional Knowledge Graph Representation</i>	
Jiaang Li, Quan Wang, Yi Liu, Licheng Zhang and Zhendong Mao	2917
<i>Exploring All-In-One Knowledge Distillation Framework for Neural Machine Translation</i>	
Zhongjian Miao, Wen Zhang, Jinsong Su, Xiang Li, Jian Luan, Yidong Chen, Bin Wang and Min Zhang	2929
<i>HistAlign: Improving Context Dependency in Language Generation by Aligning with History</i>	
David Wan, Shiyue Zhang and Mohit Bansal	2941
<i>CombLM: Adapting Black-Box Language Models through Small Fine-Tuned Models</i>	
Aitor Ormazabal, Mikel Artetxe and Eneko Agirre	2961

<i>Image Manipulation via Multi-Hop Instructions - A New Dataset and Weakly-Supervised Neuro-Symbolic Approach</i>	
Harman Singh, Poorva Garg, Mohit Gupta, Kevin Shah, Ashish Goswami, Satyam Modi, Arnab Kumar Mondal, Dinesh Khandelwal, Dinesh Garg and Parag Singla	2975
<i>Generative Spoken Language Model based on continuous word-sized audio tokens</i>	
Robin Jonathan Algayres, Yossi Adi, Tu Anh Nguyen, Jade Copet, Gabriel Synnaeve, Benoît Sagot and Emmanuel Dupoux	3008
<i>Enhancing Chat Language Models by Scaling High-quality Instructional Conversations</i>	
Ning Ding, Yulin Chen, Bokai Xu, Yujia Qin, Shengding Hu, Zhiyuan Liu, Maosong Sun and Bowen Zhou	3029
<i>Weakly-Supervised Learning of Visual Relations in Multimodal Pretraining</i>	
Emanuele Bugliarello, Aida Nematzadeh and Lisa Anne Hendricks	3052
<i>Unsupervised Grammatical Error Correction Rivaling Supervised Methods</i>	
Hannan Cao, Liping Yuan, Yuchen Zhang and Hwee Tou Ng	3072
<i>S2abEL: A Dataset for Entity Linking from Scientific Tables</i>	
Yuze Lou, Bailey Kuehl, Erin Bransom, Sergey Feldman, Aakanksha Naik and Doug Downey	3089
<i>API-Bank: A Comprehensive Benchmark for Tool-Augmented LLMs</i>	
Minghao Li, Yingxiu Zhao, Bowen Yu, Feifan Song, Hangyu Li, Haiyang Yu, Zhoujun Li, Fei Huang and Yongbin Li	3102
<i>Language and Mental Health: Measures of Emotion Dynamics from Text as Linguistic Biosocial Markers</i>	
Daniela Teodorescu, Tiffany Cheng, Alona Fyshe and Saif M. Mohammad	3117
<i>Lion: Adversarial Distillation of Proprietary Large Language Models</i>	
Yuxin Jiang, Chunkit Chan, Mingyang Chen and Wei Wang	3134
<i>Evaluating Large Language Models on Controlled Generation Tasks</i>	
Jiao Sun, Yufei Tian, Wangchunshu Zhou, Nan Xu, Qian Hu, Rahul Gupta, John Frederick Wie-ting, Nanyun Peng and Xuezhe Ma	3155
<i>DeSIQ: Towards an Unbiased, Challenging Benchmark for Social Intelligence Understanding</i>	
Xiao-Yu Guo, Yuan-Fang Li and Reza Haf	3169
<i>Why LLMs Hallucinate, and How to Get (Evidential) Closure: Perceptual, Intensional, and Extensional Learning for Faithful Natural Language Generation</i>	
Adam Bouyamourn	3181
<i>A Question Answering Framework for Decontextualizing User-facing Snippets from Scientific Documents</i>	
Benjamin Newman, Luca Soldaini, Raymond Fok, Arman Cohan and Kyle Lo	3194
<i>SLOG: A Structural Generalization Benchmark for Semantic Parsing</i>	
Bingzhi Li, Lucia Donatelli, Alexander Koller, Tal Linzen, Yuekun Yao and Najoung Kim .	3213
<i>Pushdown Layers: Encoding Recursive Structure in Transformer Language Models</i>	
Shikhar Murty, Pratyusha Sharma, Jacob Andreas and Christopher D Manning	3233
<i>Can LLMs Facilitate Interpretation of Pre-trained Language Models?</i>	
Basel Mousi, Nadir Durrani and Fahim Dalvi	3248

<i>Enhancing Low-resource Fine-grained Named Entity Recognition by Leveraging Coarse-grained Datasets</i>	
Su Ah Lee, Seokjin Oh and Woohwan Jung	3269
<i>Oolong: Investigating What Makes Transfer Learning Hard with Controlled Studies</i>	
Zhengxuan Wu, Alex Tamkin and Isabel Papadimitriou	3280
<i>Non-Autoregressive Math Word Problem Solver with Unified Tree Structure</i>	
Yi Bin, Mengqun Han, Wenhao Shi, Lei Wang, Yang Yang, See-Kiong Ng and Heng Tao Shen	3290
<i>Improving Chinese Pop Song and Hokkien Gezi Opera Singing Voice Synthesis by Enhancing Local Modeling</i>	
Peng Bai, Yue Zhou, Meizhen Zheng, Wujin Sun and Xiaodong Shi	3302
<i>What Else Do I Need to Know? The Effect of Background Information on Users' Reliance on QA Systems</i>	
Navita Goyal, Eleftheria Briakou, Amanda Stephanie Liu, Connor Baumler, Claire Bonial, Jeffrey Micher, Clare R. Voss, Marine Carpuat and Hal Daumé III	3313
<i>GROOVIST: A Metric for Grounding Objects in Visual Storytelling</i>	
Aditya Kaushik Surikuchi, Sandro Pezzelle and Raquel Fernández	3331
<i>VIBE: Topic-Driven Temporal Adaptation for Twitter Classification</i>	
Yuji Zhang, Jing Li and Wenjie Li	3340
<i>TOD-Flow: Modeling the Structure of Task-Oriented Dialogues</i>	
Sungryull Sohn, Yiwei Lyu, Anthony Zhe Liu, Lajanugen Logeswaran, Dong-Ki Kim, Dongsub Shim and Honglak Lee	3355
<i>TopWORDS-Poetry: Simultaneous Text Segmentation and Word Discovery for Classical Chinese Poetry via Bayesian Inference</i>	
Changzai Pan, Feiyue Li and Ke Deng	3372
<i>Knowledge Rumination for Pre-trained Language Models</i>	
Yunzhi Yao, Peng Wang, Shengyu Mao, Chuanqi Tan, Fei Huang, Huajun Chen and Ningyu Zhang	3387
<i>Struct-XLM: A Structure Discovery Multilingual Language Model for Enhancing Cross-lingual Transfer through Reinforcement Learning</i>	
Linjuan Wu and Weiming Lu	3405
<i>AdaSent: Efficient Domain-Adapted Sentence Embeddings for Few-Shot Classification</i>	
Yongxin Huang, Kexin Wang, Sourav Dutta, Raj Nath Patel, Goran Glavaš and Iryna Gurevych	3420
<i>Interview Evaluation: A Novel Approach for Automatic Evaluation of Conversational Question Answering Models</i>	
Xibo Li, Bowei Zou, Yifan Fan, Yanling Li, Ai Ti Aw and Yu Hong	3435
<i>TCFLE-8: a Corpus of Learner Written Productions for French as a Foreign Language and its Application to Automated Essay Scoring</i>	
Rodrigo Wilkens, Alice Pintard, David Alfter, Vincent Folny and Thomas François	3447
<i>Dancing Between Success and Failure: Edit-level Simplification Evaluation using SALSA</i>	
David Heineman, Yao Dou, Mounica Maddela and Wei Xu	3466

<i>Confidence-based Ensembling of Perspective-aware Models</i>	
Silvia Casola, Soda Marem Lo, Valerio Basile, Simona Frenda, Alessandra Teresa Cignarella, Viviana Patti and Cristina Bosco	3496
<i>ToViLaG: Your Visual-Language Generative Model is Also An Evildoer</i>	
Xinpeng Wang, Xiaoyuan Yi, Han Jiang, Shanlin Zhou, Zhihua Wei and Xing Xie	3508
<i>GPT-RE: In-context Learning for Relation Extraction using Large Language Models</i>	
Zhen Wan, Fei Cheng, Zhuoyuan Mao, Qianying Liu, Haiyue Song, Jiwei Li and Sadao Kurohashi	3534
<i>Sociocultural Norm Similarities and Differences via Situational Alignment and Explainable Textual Entailment</i>	
Sky CH-Wang, Arkadiy Saakyan, Oliver Li, Zhou Yu and Smaranda Muresan	3548
<i>INFORM : Information eNtropy based multi-step reasoning FOR large language Models</i>	
Chuyue Zhou, Wangjie You, Juntao Li, Jing Ye, Kehai Chen and Min Zhang	3565
<i>Adaptive Gating in Mixture-of-Experts based Language Models</i>	
Jiamin Li, Qiang Su, Yitao Yang, Yimin Jiang, Cong Wang and Hong Xu	3577
<i>On the Automatic Generation and Simplification of Children's Stories</i>	
Maria Valentini, Jennifer Weber, Jesus Salcido, Téa Wright, Eliana Colunga and Katharina von der Wense	3588
<i>When Do Decompositions Help for Machine Reading?</i>	
Kangda Wei, Dawn Lawrie, Benjamin Van Durme, Yunmo Chen and Orion Weller	3599
<i>The Curious Case of Hallucinatory (Un)answerability: Finding Truths in the Hidden States of Over-Confident Large Language Models</i>	
Aviv Slobodkin, Omer Goldman, Avi Caciularu, Ido Dagan and Shauli Ravfogel	3607
<i>Identifying Informational Sources in News Articles</i>	
Alexander Spangher, Nanyun Peng, Emilio Ferrara and Jonathan May	3626
<i>Retrofitting Light-weight Language Models for Emotions using Supervised Contrastive Learning</i>	
Sapan Shah, Sreedhar Reddy and Pushpak Bhattacharyya	3640
<i>Longtriever: a Pre-trained Long Text Encoder for Dense Document Retrieval</i>	
Junhan Yang, Zheng Liu, Chaozhao Li, Guangzhong Sun and Xing Xie	3655
<i>Revisiting De-Identification of Electronic Medical Records: Evaluation of Within- and Cross-Hospital Generalization</i>	
Yiyang Liu, Jinpeng Li and Enwei Zhu	3666
<i>Small Language Models Fine-tuned to Coordinate Larger Language Models improve Complex Reasoning</i>	
Gurusha Juneja, Subhabrata Dutta, Soumen Chakrabarti, Sunny Manchanda and Tanmoy Chakraborty	3675
<i>Language Representation Projection: Can We Transfer Factual Knowledge across Languages in Multilingual Language Models?</i>	
Shaoyang Xu, Junzhuo Li and Deyi Xiong	3692
<i>Structural Priming Demonstrates Abstract Grammatical Representations in Multilingual Language Models</i>	
James Michaelov, Catherine Arnett, Tyler A. Chang and Ben Bergen	3703

<i>ReasoningLM: Enabling Structural Subgraph Reasoning in Pre-trained Language Models for Question Answering over Knowledge Graph</i>	
Jinhao Jiang, Kun Zhou, Xin Zhao, Yaliang Li and Ji-Rong Wen	3721
<i>Deep Natural Language Feature Learning for Interpretable Prediction</i>	
Felipe Urrutia, Cristian Buc Calderon and Valentin Barriere	3736
<i>ROBBIE: Robust Bias Evaluation of Large Generative Language Models</i>	
David Esiobu, Xiaoqing Tan, Saghar Hosseini, Megan Ung, Yuchen Zhang, Jude Fernandes, Jane Dwivedi-Yu, Eleonora Presani, Adina Williams and Eric Michael Smith	3764
<i>Enhancing Task-oriented Dialogue Systems with Generative Post-processing Networks</i>	
Atsumoto Ohashi and Ryuichiro Higashinaka	3815
<i>Adapting Language Models to Compress Contexts</i>	
Alexis Chevalier, Alexander Wettig, Anirudh Ajith and Danqi Chen	3829
<i>Selective Labeling: How to Radically Lower Data-Labeling Costs for Document Extraction Models</i>	
Yichao Zhou, James Bradley Wendt, Navneet Potti, Jing Xie and Sandeep Tata	3847
<i>TRAVEL: Tag-Aware Conversational FAQ Retrieval via Reinforcement Learning</i>	
Yue Chen, Dingnan Jin, Chen Huang, Jia Liu and Wenqiang Lei	3861
<i>Continual Dialogue State Tracking via Example-Guided Question Answering</i>	
Hyundong Justin Cho, Andrea Madotto, Zhaojiang Lin, Khyathi Chandu, Satwik Kottur, Jing Xu, Jonathan May and Chinnadhurai Sankar	3873
<i>Lost in Translation, Found in Spans: Identifying Claims in Multilingual Social Media</i>	
Shubham Mittal, Megha Sundriyal and Preslav Nakov	3887
<i>COVID-19 Vaccine Misinformation in Middle Income Countries</i>	
Jongin Kim, Byeol Rhee Bak, Aditya Agrawal, Jiaxi Wu, Veronika J. Wirtz, Traci Hong and Derry Wijaya	3903
<i>Contrastive Learning of Sentence Embeddings from Scratch</i>	
Junlei Zhang, Zhenzhong Lan and Junxian He	3916
<i>A Rose by Any Other Name would not Smell as Sweet: Social Bias in Names Mistranslation</i>	
Sandra Camille Sandoval, Jieyu Zhao, Marine Carpuat and Hal Daumé III	3933
<i>Investigating Efficiently Extending Transformers for Long Input Summarization</i>	
Jason Phang, Yao Zhao and Peter J Liu	3946
<i>CS2W: A Chinese Spoken-to-Written Style Conversion Dataset with Multiple Conversion Types</i>	
Zishan Guo, Linhao Yu, Minghui Xu, Renren Jin and Deyi Xiong	3962
<i>Unifying Cross-Lingual Transfer across Scenarios of Resource Scarcity</i>	
Alan Ansell, Marinela Parović, Ivan Vulić, Anna Korhonen and Edoardo Ponti	3980
<i>A Tale of Pronouns: Interpretability Informs Gender Bias Mitigation for Fairer Instruction-Tuned Machine Translation</i>	
Giuseppe Attanasio, Flor Miriam Plaza del Arco, Debora Nozza and Anne Lauscher	3996
<i>DisCo: Distilled Student Models Co-training for Semi-supervised Text Mining</i>	
Weifeng Jiang, Qianren Mao, Chenghua Lin, Jianxin Li, Ting Deng, Weiye Yang and Zheng Wang	4015

<i>Dynosaur: A Dynamic Growth Paradigm for Instruction-Tuning Data Curation</i>	
Da Yin, Xiao Liu, Fan Yin, Ming Zhong, Hritik Bansal, Jiawei Han and Kai-Wei Chang . . .	4031
<i>Are All Steps Equally Important? Benchmarking Essentiality Detection in Event Processes</i>	
Haoyu Wang, Hongming Zhang, Yueguan Wang, Yuqian Deng, Muhao Chen and Dan Roth	4048
<i>Language Model is Suitable for Correction of Handwritten Mathematical Expressions Recognition</i>	
Zui Chen, Jiaqi Han, Chaofan Yang and Yi Zhou	4057
<i>Vicinal Risk Minimization for Few-Shot Cross-lingual Transfer in Abusive Language Detection</i>	
Gretel Liz De la Peña Sarraacén, Paolo Rosso, Robert Litschko, Goran Glavaš and Simone Paolo Ponzetto	4069
<i>SuperDialseg: A Large-scale Dataset for Supervised Dialogue Segmentation</i>	
Junfeng Jiang, Chengzhang Dong, Sadao Kurohashi and Akiko Aizawa	4086
<i>ATFormer: A Learned Performance Model with Transfer Learning Across Devices for Deep Learning Tensor Programs</i>	
Yang Bai, Wenqian Zhao, Shuo Yin, Zixiao Wang and Bei Yu	4102
<i>mRedditSum: A Multimodal Abstractive Summarization Dataset of Reddit Threads with Images</i>	
Keighley Overbay, Jaewoo Ahn, Fatemeh Pesaran zadeh, Joonsuk Park and Gunhee Kim . . .	4117
<i>Sparse Low-rank Adaptation of Pre-trained Language Models</i>	
Ning Ding, Xingtai Lv, Qiaosen Wang, Yulin Chen, Bowen Zhou, Zhiyuan Liu and Maosong Sun	4133
<i>Human Learning by Model Feedback: The Dynamics of Iterative Prompting with Midjourney</i>	
Shachar Don-Yehiya, Leshem Choshen and Omri Abend	4146
<i>ULF: Unsupervised Labeling Function Correction using Cross-Validation for Weak Supervision</i>	
Anastasiia Sedova and Benjamin Roth	4162
<i>The Art of SOCRATIC QUESTIONING: Recursive Thinking with Large Language Models</i>	
Jingyuan Qi, Zhiyang Xu, Ying Shen, Minqian Liu, Di Jin, Qifan Wang and Lifu Huang . . .	4177
<i>Ideology Takes Multiple Looks: A High-Quality Dataset for Multifaceted Ideology Detection</i>	
Songtao Liu, Ziling Luo, Minghua Xu, Lixiao Wei, Ziyao Wei, Han Yu, Wei Xiang and Bang Wang	4200
<i>Transductive Learning for Textual Few-Shot Classification in API-based Embedding Models</i>	
Pierre Colombo, Victor Pellegrain, Malik Boudiaf, Myriam Tami, Victor Storch, Ismail Ben Ayed and Pablo Piantanida	4214
<i>MEGA: Multilingual Evaluation of Generative AI</i>	
Kabir Ahuja, Harshita Diddee, Rishav Hada, Millicent Ochieng, Krithika Ramesh, Prachi Jain, Akshay Nambi, Tanuja Ganu, Sameer Segal, Mohamed Ahmed, Kalika Bali and Sunayana Sitaram	4232
<i>Support or Refute: Analyzing the Stance of Evidence to Detect Out-of-Context Mis- and Disinformation</i>	
Xin Yuan, Jie Guo, Weidong Qiu, Zheng Huang and Shujun Li	4268
<i>Video-Helpful Multimodal Machine Translation</i>	
Yihang Li, Shuichiro Shimizu, Chenhui Chu, Sadao Kurohashi and Wei Li	4281
<i>Large Language Models are Temporal and Causal Reasoners for Video Question Answering</i>	
Dohwan Ko, Ji Soo Lee, Woo-Young Kang, Byungseok Roh and Hyunwoo J. Kim	4300

<i>Uncertainty Guided Global Memory Improves Multi-Hop Question Answering</i>	
Alsu Sagirova and Mikhail Burtsev	4317
<i>Prompting Large Language Models with Chain-of-Thought for Few-Shot Knowledge Base Question Generation</i>	
Yuanyuan Liang, Jianing Wang, Hanlun Zhu, Lei Wang, Weining Qian and Yunshi Lan	4329
<i>TrojanSQL: SQL Injection against Natural Language Interface to Database</i>	
Jinchuan Zhang, Yan Zhou, Binyuan Hui, Yaxin Liu, Ziming Li and Songlin Hu	4344
<i>Preserving Privacy Through Dememorization: An Unlearning Technique For Mitigating Memorization Risks In Language Models</i>	
Aly M. Kassem, Omar Mahmoud and Sherif Saad	4360
<i>MingOfficial: A Ming Official Career Dataset and a Historical Context-Aware Representation Learning Framework</i>	
You-Jun Chen, Hsin-Yi Hsieh, Yu Tung Lin, Yingtao Tian, Bert Chan, Yu-Sin Liu, Yi-Hsuan Lin and Richard Tzong-Han Tsai	4380
<i>DPP-TTS: Diversifying prosodic features of speech via determinantal point processes</i>	
Seongho Joo, Hyukhun Koh and Kyomin Jung	4402
<i>Meta-Learning Online Adaptation of Language Models</i>	
Nathan Zixia Hu, Eric Mitchell, Christopher D Manning and Chelsea Finn	4418
<i>Self-Detoxifying Language Models via Toxification Reversal</i>	
Chak Tou Leong, Yi Cheng, Jiashuo Wang, Jian Wang and Wenjie Li	4433
<i>Interactive Text Generation</i>	
Felix Faltings, Michel Galley, Kianté Brantley, Baolin Peng, Weixin Cai, Yizhe Zhang, Jianfeng Gao and Bill Dolan	4450
<i>Knowledge Distillation $\approx$ Label Smoothing: Fact or Fallacy?</i>	
Md Arafat Sultan	4469
<i>Analyzing Cognitive Plausibility of Subword Tokenization</i>	
Lisa Beinborn and Yuval Pinter	4478
<i>POE: Process of Elimination for Multiple Choice Reasoning</i>	
Chenkai Ma and Xinya Du	4487
<i>NeuSTIP: A Neuro-Symbolic Model for Link and Time Prediction in Temporal Knowledge Graphs</i>	
Ishaan Singh, Navdeep Kaur, Garima Gaur and Mausam	4497
<i>Standardizing Distress Analysis: Emotion-Driven Distress Identification and Cause Extraction (DICE) in Multimodal Online Posts</i>	
Gopendra Vikram Singh, Soumitra Ghosh, Atul Verma, Chetna Painkra and Asif Ekbal	4517
<i>Out-of-Distribution Generalization in Natural Language Processing: Past, Present, and Future</i>	
Linyi Yang, Yaoxian Song, Xuan Ren, Chenyang Lyu, Yidong Wang, Jingming Zhuo, Lingqiao Liu, Jindong Wang, Jennifer Foster and Yue Zhang	4533
<i>Noisy Exemplars Make Large Language Models More Robust: A Domain-Agnostic Behavioral Analysis</i>	
Hongyi Zheng and Abulhair Saparov	4560
<i>Can Large Language Models Capture Dissenting Human Voices?</i>	
Noah Lee, Na Min An and James Thorne	4569

<i>DecoMT: Decomposed Prompting for Machine Translation Between Related Languages using Large Language Models</i>	
Ratish Puduppully, Anoop Kunchukuttan, Raj Dabre, Ai Ti Aw and Nancy F. Chen	4586
<i>Prototype-based HyperAdapter for Sample-Efficient Multi-task Tuning</i>	
Hao Zhao, Jie Fu and Zhaofeng He	4603
<i>Towards Building More Robust NER datasets: An Empirical Study on NER Dataset Bias from a Dataset Difficulty View</i>	
Ruotian Ma, Xiaolei Wang, Xin Zhou, Qi Zhang and Xuanjing Huang	4616
<i>GradSim: Gradient-Based Language Grouping for Effective Multilingual Training</i>	
Mingyang Wang, Heike Adel, Lukas Lange, Jannik Strötgen and Hinrich Schuetze	4631
<i>Discovering Universal Geometry in Embeddings with ICA</i>	
Hiroaki Yamagiwa, Momose Oyama and Hidetoshi Shimodaira	4647
<i>Toward a Critical Toponymy Framework for Named Entity Recognition: A Case Study of Airbnb in New York City</i>	
Mikael Brunila, Jack LaViolette, Sky CH-Wang, Priyanka Verma, Clara Féré and Grant McKenzie	4676
<i>Well Begun is Half Done: Generator-agnostic Knowledge Pre-Selection for Knowledge-Grounded Dialogue</i>	
Lang Qin, Yao Zhang, Hongru Liang, Jun Wang and Zhenglu Yang	4696
<i>Merging Generated and Retrieved Knowledge for Open-Domain QA</i>	
Yunxiang Zhang, Muhammad Khalifa, Lajanugen Logeswaran, Moontae Lee, Honglak Lee and Lu Wang	4710
<i>Best of Both Worlds: Towards Improving Temporal Knowledge Base Question Answering via Targeted Fact Extraction</i>	
Nithish Kannan, Udit Sharma, Sumit Neelam, Dinesh Khandelwal, Shajith Ikkal, Hima Karanam and L Venkata Subramaniam	4729
<i>Text Fact Transfer</i>	
Nishant Balepur, Jie Huang and Kevin Chang	4745
<i>A Cheaper and Better Diffusion Language Model with Soft-Masked Noise</i>	
Jiaao Chen, Aston Zhang, Mu Li, Alex Smola and Diyi Yang	4765
<i>Mirages. On Anthropomorphism in Dialogue Systems</i>	
Gavin Abercrombie, Amanda Cercas Curry, Tanvi Dinkar, Verena Rieser and Zeerak Talat	4776
<i>Cognitive Dissonance: Why Do Language Model Outputs Disagree with Internal Representations of Truthfulness?</i>	
Kevin Liu, Stephen Casper, Dylan Hadfield-Menell and Jacob Andreas	4791
<i>KEBAP: Korean Error Explainable Benchmark Dataset for ASR and Post-processing</i>	
Seonmin Koo, Chanjun Park, Jinsung Kim, Jaehyung Seo, Sugyeong Eo, Hyeonseok Moon and Heuiseok Lim	4798
<i>Adaptive Policy with Wait-k Model for Simultaneous Translation</i>	
Libo Zhao, Kai Fan, Wei Luo, Wu Jing, Shushu Wang, Ziqian Zeng and Zhongqiang Huang	4816
<i>Cross-Document Event Coreference Resolution on Discourse Structure</i>	
Xinyu Chen, Sheng Xu, Peifeng Li and Qiaoming Zhu	4833

<i>Post-hoc Utterance Refining Method by Entity Mining for Faithful Knowledge Grounded Conversations</i> Yoonna Jang, Suhyune Son, Jeongwoo Lee, Junyoung Son, Yuna Hur, Jungwoo Lim, Hyeonseok Moon, Kisu Yang and Heuseok Lim	4844
<i>Can We Edit Factual Knowledge by In-Context Learning?</i> Ce Zheng, Lei Li, Qingxiu Dong, Yuxuan Fan, Zhiyong Wu, Jingjing Xu and Baobao Chang	4862
<i>EDIS: Entity-Driven Image Search over Multimodal Web Content</i> Siqi Liu, Weixi Feng, Tsu-Jui Fu, Wenhua Chen and William Yang Wang	4877
<i>GQA: Training Generalized Multi-Query Transformer Models from Multi-Head Checkpoints</i> Joshua Ainslie, James Lee-Thorp, Michiel de Jong, Yury Zemlyanskiy, Federico Lebron and Sumit Sanghai	4895
<i>Towards a Mechanistic Interpretation of Multi-Step Reasoning Capabilities of Language Models</i> Yifan Hou, Jiaoda Li, Yu Fei, Alessandro Stolfo, Wangchunshu Zhou, Guangtao Zeng, Antoine Bosselut and Mrinmaya Sachan	4902
<i>BiasX: “Thinking Slow” in Toxic Content Moderation with Explanations of Implied Social Biases</i> Yiming Zhang, Sravani Uttara Nanduri, Liwei Jiang, Tongshuang Wu and Maarten Sap	4920
<i>Text encoders bottleneck compositionality in contrastive vision-language models</i> Amita Kamath, Jack Hessel and Kai-Wei Chang	4933
<i>Ignore This Title and HackAPrompt: Exposing Systemic Vulnerabilities of LLMs Through a Global Prompt Hacking Competition</i> Sander V Schulhoff, Jeremy Pinto, Ansum Khan, Louis-François Bouchard, Chenglei Si, Svetlana Anati, Valen Tagliabue, Anson Liu Kost, Christopher R Carnahan and Jordan Lee Boyd-Graber ..	4945
<i>MMNMT: Modularizing Multilingual Neural Machine Translation with Flexibly Assembled MoE and Dense Blocks</i> Shangjie Li, Xiangpeng Wei, Shaolin Zhu, Jun Xie, Baosong Yang and Deyi Xiong	4978
<i>Localizing Active Objects from Egocentric Vision with Symbolic World Knowledge</i> Te-Lin Wu, Yu Zhou and Nanyun Peng	4991
<i>Introducing Rhetorical Parallelism Detection: A New Task with Datasets, Metrics, and Baselines</i> Stephen Bothwell, Justin DeBenedetto, Theresa Crnkovich, Hildegund Muller and David Chiang	5007
<i>Prompting is not a substitute for probability measurements in large language models</i> Jennifer Hu and Roger P. Levy	5040
<i>Parameter-Efficient Language Model Tuning with Active Learning in Low-Resource Settings</i> Josip Jukić and Jan Snajder	5061
<i>Stop Uploading Test Data in Plain Text: Practical Strategies for Mitigating Data Contamination by Evaluation Benchmarks</i> Alon Jacovi, Avi Caciularu, Omer Goldman and Yoav Goldberg	5075
<i>CoLT5: Faster Long-Range Transformers with Conditional Computation</i> Joshua Ainslie, Tao Lei, Michiel de Jong, Santiago Ontanon, Siddhartha Brahma, Yury Zemlyanskiy, David Uthus, Mandy Guo, James Lee-Thorp, Yi Tay, Yun-Hsuan Sung and Sumit Sanghai	5085
<i>DiSTRICK: Dialogue State Tracking with Retriever Driven In-Context Tuning</i> Praveen Venkateswaran, Evelyn Duesterwald and Vatche Isahagian	5101

<i>Cross-Cultural Analysis of Human Values, Morals, and Biases in Folk Tales</i>	
Winston Wu, Lu Wang and Rada Mihalcea	5113
<i>Non-Programmers Can Label Programs Indirectly via Active Examples: A Case Study with Text-to-SQL</i>	
Ruiqi Zhong, Charlie Victor Snell, Dan Klein and Jason Eisner	5126
<i>LINC: A Neurosymbolic Approach for Logical Reasoning by Combining Language Models with First-Order Logic Provers</i>	
Theo X. Olausson, Alex Gu, Ben Lipkin, Cedegao E. Zhang, Armando Solar-Lezama, Joshua B. Tenenbaum and Roger P. Levy	5153
<i>Non-autoregressive Streaming Transformer for Simultaneous Translation</i>	
Zhengrui Ma, Shaolei Zhang, Shoutao Guo, Chenze Shao, Min Zhang and Yang Feng	5177
<i>ViSoBERT: A Pre-Trained Language Model for Vietnamese Social Media Text Processing</i>	
Nam Quoc Nguyen, Thang Chau Phan, Duc-Vu Nguyen and Kiet Van Nguyen	5191
<i>RAPL: A Relation-Aware Prototype Learning Approach for Few-Shot Document-Level Relation Extraction</i>	
Shiao Meng, Xuming Hu, Aiwei Liu, Shuang Li, Fukun Ma, Yawen Yang and Lijie Wen	5208
<i>GeoLM: Empowering Language Models for Geospatially Grounded Language Understanding</i>	
Zekun Li, Wenxuan Zhou, Yao-Yi Chiang and Muhao Chen	5227
<i>Cross-Modal Conceptualization in Bottleneck Models</i>	
Danis Alukaev, Semen Kiselev, Ilya Pershin, Bulat Ibragimov, Vladimir V. Ivanov, Alexey Kor-naev and Ivan Titov	5241
<i>LLM-Adapters: An Adapter Family for Parameter-Efficient Fine-Tuning of Large Language Models</i>	
Zhiqiang Hu, Lei Wang, Yihuai Lan, Wanyu Xu, Ee-Peng Lim, Lidong Bing, Xing Xu, Soujanya Poria and Roy Ka-Wei Lee	5254
<i>DREAM: Deployment of Recombination and Ensembles in Argument Mining</i>	
Florian Ruosch, Cristina Sarasua and Abraham Bernstein	5277
<i>MILDSum: A Novel Benchmark Dataset for Multilingual Summarization of Indian Legal Case Judgments</i>	
Debtanu Datta, Shubham Soni, Rajdeep Mukherjee and Saptarshi Ghosh	5291
<i>Query Rewriting in Retrieval-Augmented Large Language Models</i>	
Xinbei Ma, Yeyun Gong, Pengcheng He, Hai Zhao and Nan Duan	5303
<i>PromptMix: A Class Boundary Augmentation Method for Large Language Model Distillation</i>	
Gaurav Sahu, Olga Vechtomova, Dzmitry Bahdanau and Issam H. Laradji	5316
<i>COHESENTIA: A Novel Benchmark of Incremental versus Holistic Assessment of Coherence in Generated Texts</i>	
Aviya Maimon and Reut Tsarfaty	5328
<i>QUDeval: The Evaluation of Questions Under Discussion Discourse Parsing</i>	
Yating Wu, Ritika Rajesh Mangla, Greg Durrett and Junyi Jessy Li	5344
<i>PRCA: Fitting Black-Box Large Language Models for Retrieval Question Answering via Pluggable Reward-Driven Contextual Adapter</i>	
Haoyan Yang, Zhitao Li, Yong Zhang, Jianzong Wang, Ning Cheng, Ming Li and Jing Xiao	5364
<i>Exploring Chain of Thought Style Prompting for Text-to-SQL</i>	
Chang-Yu Tai, Ziru Chen, Tianshu Zhang, Xiang Deng and Huan Sun	5376

<i>Efficient Algorithms for Recognizing Weighted Tree-Adjoining Languages</i> Alexandra Butoi, Tim Vieira, Ryan Cotterell and David Chiang	5394
<i>Harnessing Black-Box Control to Boost Commonsense in LM's Generation</i> Yufei Tian, Felix Zhang and Nanyun Peng	5417
<i>Just Ask for Calibration: Strategies for Eliciting Calibrated Confidence Scores from Language Models Fine-Tuned with Human Feedback</i> Katherine Tian, Eric Mitchell, Allan Zhou, Archit Sharma, Rafael Rafailov, Huaxiu Yao, Chelsea Finn and Christopher D Manning	5433
<i>Representative Demonstration Selection for In-Context Learning with Two-Stage Determinantal Point Process</i> Zhao Yang, Yuanzhe Zhang, Dianbo Sui, Cao Liu, Jun Zhao and Kang Liu	5443
<i>The Effect of Scaling, Retrieval Augmentation and Form on the Factual Consistency of Language Models</i> Lovisa Hagström, Denitsa Saynova, Tobias Norlund, Moa Johansson and Richard Johansson	5457
<i>ViPE: Visualise Pretty-much Everything</i> Hassan Shahmohammadi, Adhiraj Ghosh and Hendrik Lensch	5477
<i>Semi-automatic Data Enhancement for Document-Level Relation Extraction with Distant Supervision from Large Language Models</i> Junpeng Li, Zixia Jia and Zilong Zheng	5495
<i>Navigating the Grey Area: How Expressions of Uncertainty and Overconfidence Affect Language Models</i> Kaitlyn Zhou, Dan Jurafsky and Tatsunori Hashimoto	5506
<i>Elaborative Simplification as Implicit Questions Under Discussion</i> Yating Wu, William Berkeley Sheffield, Kyle Mahowald and Junyi Jessy Li	5525
<i>EntSUMv2: Dataset, Models and Evaluation for More Abstractive Entity-Centric Summarization</i> Dhruv Mehra, Lingjue Xie, Ella Hofmann-Coyle, Mayank Kulkarni and Daniel Preotiuc-Pietro	5538
<i>SciRepEval: A Multi-Format Benchmark for Scientific Document Representations</i> Amanpreet Singh, Mike D'Arcy, Arman Cohan, Doug Downey and Sergey Feldman	5548
<i>A Diachronic Perspective on User Trust in AI under Uncertainty</i> Shehzaad Zuzar Dhuliawala, Vilém Zouhar, Mennatallah El-Assady and Mrinmaya Sachan	5567
<i>CT-GAT: Cross-Task Generative Adversarial Attack based on Transferability</i> Minxuan Lv, Chengwei Dai, Kun Li, Wei Zhou and Songlin Hu	5581
<i>Improving Long Document Topic Segmentation Models With Enhanced Coherence Modeling</i> Hai Yu, Chong Deng, Qinglin Zhang, Jiaqing Liu, Qian Chen and Wen Wang	5592
<i>Dialogue Chain-of-Thought Distillation for Commonsense-aware Conversational Agents</i> Hyunjoo Chae, Yongho Song, Kai Tzu-iunn Ong, Taeyoon Kwon, Minjin Kim, Youngjae Yu, Dongha Lee, Dongyeop Kang and Jinyoung Yeo	5606
<i>Information Value: Measuring Utterance Predictability as Distance from Plausible Alternatives</i> Mario Giulianelli, Sarenne Wallbridge and Raquel Fernández	5633
<i>Generating Commonsense Counterfactuals for Stable Relation Extraction</i> Xin Miao, Yongqi Li and Tieyun Qian	5654

<i>C-STs: Conditional Semantic Textual Similarity</i>	
Ameet Deshpande, Carlos E Jimenez, Howard Chen, Vishvak Murahari, Victoria Graf, Tanmay Rajpurohit, Ashwin Kalyan, Danqi Chen and Karthik R Narasimhan	5669
<i>Cross-lingual Transfer Can Worsen Bias in Sentiment Analysis</i>	
Seraphina Goldfarb-Tarrant, Björn Ross and Adam Lopez	5691
<i>Rumor Detection on Social Media with Crowd Intelligence and ChatGPT-Assisted Networks</i>	
Chang Yang, Peng Zhang, Wenbo Qiao, Hui Gao and Jiaming Zhao	5705
<i>Grounding Visual Illusions in Language: Do Vision-Language Models Perceive Illusions Like Humans?</i>	
Yichi Zhang, Jiayi Pan, Yuchen Zhou, Rui Pan and Joyce Chai	5718
<i>Analysing State-Backed Propaganda Websites: a New Dataset and Linguistic Study</i>	
Freddy Heppell, Kalina Bontcheva and Carolina Scarton	5729
<i>Controllable Contrastive Generation for Multilingual Biomedical Entity Linking</i>	
Tiantian Zhu, Yang Qin, Qingcai Chen, Xin Mu, Changlong Yu and Yang Xiang	5742
<i>HyperRouter: Towards Efficient Training and Inference of Sparse Mixture of Experts</i>	
Truong Giang Do, Le Huy Khiem, Quang Pham, TrungTin Nguyen, Thanh-Nam Doan, Binh T. Nguyen, Chenghao Liu, Savitha Ramasamy, Xiaoli Li and Steven Hoi	5754
<i>MediaHG: Rethinking Eye-catching Features in Social Media Headline Generation</i>	
Boning Zhang and Yang Yang	5766
<i>Fine-tuned LLMs Know More, Hallucinate Less with Few-Shot Sequence-to-Sequence Semantic Parsing over Wikidata</i>	
Silei Xu, Shicheng Liu, Theo Culhane, Elizaveta Pertseva, Meng-Hsi Wu, Sina Semnani and Monica Lam	5778
<i>ZEROTOP: Zero-Shot Task-Oriented Semantic Parsing using Large Language Models</i>	
Dheeraj Mekala, Jason Andrew Wolfe and Subhro Roy	5792
<i>Efficient Grammatical Error Correction Via Multi-Task Training and Optimized Training Schedule</i>	
Andrey Bout, Alexander Podolskiy, Sergey Nikolenko and Irina Piontkovskaya	5800
<i>The BLA Benchmark: Investigating Basic Language Abilities of Pre-Trained Multimodal Models</i>	
Xinyi Chen, Raquel Fernández and Sandro Pezzelle	5817
<i>RainProof: An Umbrella to Shield Text Generator from Out-Of-Distribution Data</i>	
Maxime Darrin, Pablo Piantanida and Pierre Colombo	5831
<i>KEPL: Knowledge Enhanced Prompt Learning for Chinese Hypernym-Hyponym Extraction</i>	
Ningchen Ma, Dong Wang, Hongyun Bao, Lei He and Suncong Zheng	5858
<i>Ditto: A Simple and Efficient Approach to Improve Sentence Embeddings</i>	
Qian Chen, Wen Wang, Qinglin Zhang, Siqi Zheng, Chong Deng, Hai Yu, Jiaqing Liu, Yukun Ma and Chong Zhang	5868
<i>Preserving Knowledge Invariance: Rethinking Robustness Evaluation of Open Information Extraction</i>	
Ji Qi, Chuchun Zhang, Xiaozhi Wang, Kaisheng Zeng, Jifan Yu, Jinxin Liu, Lei Hou, Juanzi Li and Xu Bin	5876
<i>Why Should This Article Be Deleted? Transparent Stance Detection in Multilingual Wikipedia Editor Discussions</i>	
Lucie-Aimée Kaffee, Arnav Arora and Isabelle Augenstein	5891

<i>Fast and Robust Early-Exiting Framework for Autoregressive Language Models with Synchronized Parallel Decoding</i>	
Sangmin Bae, Jongwoo Ko, Hwanjun Song and Se-Young Yun	5910
<i>End-to-end Task-oriented Dialogue: A Survey of Tasks, Methods, and Future Directions</i>	
Libo Qin, Wenbo Pan, Qiguang Chen, Lizi Liao, Zhou Yu, Yue Zhang, Wanxiang Che and Min Li	5925
<i>Answering Questions by Meta-Reasoning over Multiple Chains of Thought</i>	
Ori Yoran, Tomer Wolfson, Ben Bogin, Uri Katz, Daniel Deutch and Jonathan Berant	5942
<i>INSTRUCTSCORE: Towards Explainable Text Generation Evaluation with Automatic Feedback</i>	
Wenda Xu, Danqing Wang, Liangming Pan, Zhenqiao Song, Markus Freitag, William Yang Wang and Lei Li	5967
<i>Multi-level Contrastive Learning for Script-based Character Understanding</i>	
Dawei Li, Hengyuan Zhang, Yanran Li and Shiping Yang	5995
<i>CHEF in the Language Kitchen: A Generative Data Augmentation Leveraging Korean Morpheme Ingredients</i>	
Jaehyung Seo, Hyeonseok Moon, Jaewook Lee, Sugyeong Eo, Chanjun Park and Heuiseok Lim	6014
<i>Automatic Debate Evaluation with Argumentation Semantics and Natural Language Argument Graph Networks</i>	
Ramon Ruiz-Dolz, Stella Heras and Ana Garcia	6030
<i>Transfer-Free Data-Efficient Multilingual Slot Labeling</i>	
Evgeniia Razumovskaia, Ivan Vulić and Anna Korhonen	6041
<i>Towards Interpretable Mental Health Analysis with Large Language Models</i>	
Kailai Yang, Shaoxiong Ji, Tianlin Zhang, Qianqian Xie, Ziyang Kuang and Sophia Ananiadou	6056
<i>Learning to Rank Generation with Pairwise Partial Rewards</i>	
Youngwon Lee, Jinu Lee and Seung-won Hwang	6078
<i>GreedyCAS: Unsupervised Scientific Abstract Segmentation with Normalized Mutual Information</i>	
Yingqiang Gao, Jessica Lam, Nianlong Gu and Richard Hahnloser	6093
<i>Spoiler Detection as Semantic Text Matching</i>	
Ryan Tran, Canwen Xu and Julian McAuley	6109
<i>Multimodal Embodied Plan Prediction Augmented with Synthetic Embodied Dialogue</i>	
Aishwarya Padmakumar, Mert Inan, Spandana Gella, Patrick L. Lange and Dilek Hakkani-Tur	6114
<i>GEM: Gestalt Enhanced Markup Language Model for Web Understanding via Render Tree</i>	
Zirui Shao, Feiyu Gao, Zhongda Qi, Hangdi Xing, Jiajun Bu, Zhi Yu, Qi Zheng and Xiaozhong Liu	6132
<i>Abstractive Open Information Extraction</i>	
Kevin Song Pei, Ishan Jindal and Kevin Chang	6146
<i>CoSyn: Detecting Implicit Hate Speech in Online Conversations Using a Context Synergized Hyperbolic Network</i>	
Sreyan Ghosh, Manan Suri, Purva Chiniya, Utkarsh Tyagi, Sonal Kumar and Dinesh Manocha	6159

<i>CLEME: Debiasing Multi-reference Evaluation for Grammatical Error Correction</i>	
Jingheng Ye, Yinghui Li, Qingyu Zhou, Yangning Li, Shirong Ma, Hai-Tao Zheng and Ying Shen	
6174	
<i>Dynamic Top-k Estimation Consolidates Disagreement between Feature Attribution Methods</i>	
Jonathan Kamp, Lisa Beinborn and Antske Fokkens	6190
<i>SentiStream: A Co-Training Framework for Adaptive Online Sentiment Analysis in Evolving Data Streams</i>	
Yuhao Wu, Karthick Sharma, Chun Wei Seah and Shuhao Zhang	6198
<i>HyperNetwork-based Decoupling to Improve Model Generalization for Few-Shot Relation Extraction</i>	
Liang Zhang, Chulun Zhou, Fandong Meng, Jinsong Su, Yidong Chen and Jie Zhou	6213
<i>Solving Hard Analogy Questions with Relation Embedding Chains</i>	
Nitesh Kumar and Steven Schockaert	6224
<i>Modeling Empathic Similarity in Personal Narratives</i>	
Jocelyn J Shen, Maarten Sap, Pedro Colon-Hernandez, Hae Won Park and Cynthia Breazeal	6237
<i>Tree Prompting: Efficient Task Adaptation without Fine-Tuning</i>	
Chandan Singh, John Xavier Morris, Alexander M Rush, Jianfeng Gao and Yuntian Deng	6253
<i>Baize: An Open-Source Chat Model with Parameter-Efficient Tuning on Self-Chat Data</i>	
Canwen Xu, Daya Guo, Nan Duan and Julian McAuley	6268
<i>Empathy Intent Drives Empathy Detection</i>	
Liting Jiang, Di Wu, Bohui Mao, Yanbing Li and Wushour Slamu	6279
<i>Adaptive End-to-End Metric Learning for Zero-Shot Cross-Domain Slot Filling</i>	
Yuanjun Shi, Linzhi Wu and Minglai Shao	6291
<i>BasahaCorpus: An Expanded Linguistic Resource for Readability Assessment in Central Philippine Languages</i>	
Joseph Marvin Imperial and Ekaterina Kochmar	6302
<i>ReTAG: Reasoning Aware Table to Analytic Text Generation</i>	
Deepanway Ghosal, Preksha Nema and Aravindan Raghuv eer	6310
<i>Beyond Factuality: A Comprehensive Evaluation of Large Language Models as Knowledge Generators</i>	
Liang Chen, Yang Deng, Yatao Bian, Zeyu Qin, Bingzhe Wu, Tat-Seng Chua and Kam-Fai Wong	6325
<i>Compressing Context to Enhance Inference Efficiency of Large Language Models</i>	
Yucheng Li, Bo Dong, Frank Guerin and Chenghua Lin	6342
<i>MoT: Memory-of-Thought Enables ChatGPT to Self-Improve</i>	
Xiaonan Li and Xipeng Qiu	6354
<i>4 and 7-bit Labeling for Projective and Non-Projective Dependency Trees</i>	
Carlos Gómez-Rodríguez, Diego Roca and David Vilares	6375
<i>Can You Follow Me? Testing Situational Understanding for ChatGPT</i>	
Chenghao Yang and Allyson Ettinger	6385
<i>Towards Reliable Misinformation Mitigation: Generalization, Uncertainty, and GPT-4</i>	
Kellin Pelrine, Anne Imouza, Camille Thibault, Meilina Reksoprodjo, Caleb Alexander Gupta, Joel Christoph, Jean-François Godbout and Reihaneh Rabbany	6399

<i>Advancements in Arabic Grammatical Error Detection and Correction: An Empirical Investigation</i>	
Bashar Alhafni, Go Inoue, Christian Khairallah and Nizar Habash	6430
<i>HaluEval: A Large-Scale Hallucination Evaluation Benchmark for Large Language Models</i>	
Junyi Li, Xiaoxue Cheng, Xin Zhao, Jian-Yun Nie and Ji-Rong Wen	6449
<i>Enabling Large Language Models to Generate Text with Citations</i>	
Tianyu Gao, Howard Yen, Jiatong Yu and Danqi Chen	6465
<i>Revisiting Machine Translation for Cross-lingual Classification</i>	
Mikel Artetxe, Vedanuj Goswami, Shruti Bhosale, Angela Fan and Luke Zettlemoyer	6489
<i>Pre-Trained Language Models Augmented with Synthetic Scanpaths for Natural Language Understanding</i>	
Shuwen Deng, Paul Prasse, David Robert Reich, Tobias Scheffer and Lena Ann Jäger	6500
<i>Counting the Bugs in ChatGPT's Wugs: A Multilingual Investigation into the Morphological Capabilities of a Large Language Model</i>	
Leonie Weissweiler, Valentin Hofmann, Anjali Kantharuban, Anna Cai, Ritam Dutt, Amey Hengle, Anubha Kabra, Atharva Kulkarni, Abhishek Vijayakumar, Haofei Yu, Hinrich Schuetze, Kemal Oflazer and David R Mortensen	6508
<i>Adapt in Contexts: Retrieval-Augmented Domain Adaptation via In-Context Learning</i>	
Quanyu Long, Wenya Wang and Sinno Jialin Pan	6525
<i>Understanding the Inner-workings of Language Models Through Representation Dissimilarity</i>	
Davis Brown, Charles Godfrey, Nicholas Konz, Jonathan Tu and Henry Kvinge	6543
<i>Efficient Classification of Long Documents via State-Space Models</i>	
Peng Lu, Suyuchen Wang, Mehdi Rezagholizadeh, Bang Liu and Ivan Kobzyev	6559
<i>Dual-Feedback Knowledge Retrieval for Task-Oriented Dialogue Systems</i>	
Tianyuan Shi, Liangzhi Li, Zijian Lin, Tao Yang, Xiaojun Quan and Qifan Wang	6566
<i>Construction Artifacts in Metaphor Identification Datasets</i>	
Joanne Boisson, Luis Espinosa-Anke and Jose Camacho-Collados	6581
<i>MAF: Multi-Aspect Feedback for Improving Reasoning in Large Language Models</i>	
Deepak Nathani, David Wang, Liangming Pan and William Yang Wang	6591
<i>Granularity Matters: Pathological Graph-driven Cross-modal Alignment for Brain CT Report Generation</i>	
Yanzhao Shi, Junzhong Ji, Xiaodan Zhang, Liangqiong Qu and Ying Liu	6617
<i>Enhancing Structured Evidence Extraction for Fact Verification</i>	
Zirui Wu, Nan Hu and Yansong Feng	6631
<i>Rethinking Model Selection and Decoding for Keyphrase Generation with Pre-trained Sequence-to-Sequence Models</i>	
Di Wu, Wasi Uddin Ahmad and Kai-Wei Chang	6642
<i>A Fair and In-Depth Evaluation of Existing End-to-End Entity Linking Systems</i>	
Hannah Bast, Matthias Hertel and Natalie Prange	6659
<i>A Multi-Task Dataset for Assessing Discourse Coherence in Chinese Essays: Structure, Theme, and Logic Analysis</i>	
Hongyi Wu, Xinshu Shen, Man Lan, Shaoguang Mao, Xiaopeng Bai and Yuanbin Wu	6673

<i>SKD-NER: Continual Named Entity Recognition via Span-based Knowledge Distillation with Reinforcement Learning</i>	
Yi Chen and Liang He	6689
<i>Lifelong Sequence Generation with Dynamic Module Expansion and Adaptation</i>	
Chengwei Qin, Chen Chen and Shafiq Joty	6701
<i>When the Majority is Wrong: Modeling Annotator Disagreement for Subjective Tasks</i>	
Eve Fleisig, Rediet Abebe and Dan Klein	6715
<i>Lazy-k Decoding: Constrained Decoding for Information Extraction</i>	
Arthur Hemmer, Mickael Coustaty, Nicola Bartolo, Jerome Brachat and Jean-marc Ogier ..	6727
<i>Personalized Distillation: Empowering Open-Sourced LLMs with Adaptive Learning for Code Generation</i>	
Hailin Chen, Amrita Saha, Steven Hoi and Shafiq Joty	6737
<i>Do Language Models Have a Common Sense regarding Time? Revisiting Temporal Commonsense Reasoning in the Era of Large Language Models</i>	
Raghav Jain, Daivik Sojitra, Arkadeep Acharya, Sriparna Saha, Adam Jatowt and Sandipan Dandapat	6750
<i>Comparing Styles across Languages</i>	
Shreya Havaldar, Matthew Pressimone, Eric Wong and Lyle Ungar	6775
<i>Event Causality Extraction via Implicit Cause-Effect Interactions</i>	
Jintao Liu, Zequn Zhang, Kaiwen Wei, Zhi Guo, Xian Sun, Li Jin and Xiaoyu Li	6792
<i>Evaluation of African American Language Bias in Natural Language Generation</i>	
Nicholas Deas, Jessica A Grieser, Shana Kleiner, Desmond U. Patton, Elsbeth Turcan and Kathleen McKeown	6805
<i>A Systematic Study of Performance Disparities in Multilingual Task-Oriented Dialogue Systems</i>	
Songbo Hu, Han Zhou, Moy Yuan, Milan Gritta, Guchun Zhang, Ignacio Iacobacci, Anna Korhonen and Ivan Vulić	6825
<i>Cognate Transformer for Automated Phonological Reconstruction and Cognate Reflex Prediction</i>	
V.S.D.S.Mahesh Akavarapu and Arnab Bhattacharya	6852
<i>Inference-Time Policy Adapters (IPA): Tailoring Extreme-Scale LMs without Fine-tuning</i>	
Ximing Lu, Faeze Brahman, Peter West, Jaehun Jung, Khyathi Chandu, Abhilasha Ravichander, Prithviraj Ammanabrolu, Liwei Jiang, Sahana Ramnath, Nouha Dziri, Jillian Fisher, Bill Yuchen Lin, Skyler Hallinan, Lianhui Qin, Xiang Ren, Sean Welleck and Yejin Choi	6863
<i>Weakly Supervised Semantic Parsing with Execution-based Spurious Program Filtering</i>	
Kang-il Lee, Segwang Kim and Kyomin Jung	6884
<i>Taxonomy Expansion for Named Entity Recognition</i>	
Karthikeyan K, Yogarshi Vyas, Jie Ma, Giovanni Paolini, Neha Anna John, Shuai Wang, Yassine Benajiba, Vittorio Castelli, Dan Roth and Miguel Ballesteros	6895
<i>Rather a Nurse than a Physician - Contrastive Explanations under Investigation</i>	
Oliver Eberle, Ilias Chalkidis, Laura Cabello and Stephanie Brandl	6907
<i>EtiCor: Corpus for Analyzing LLMs for Etiquettes</i>	
Ashutosh Dwivedi, Pradhyumna Lavania and Ashutosh Modi	6921

<i>An Investigation of LLMs' Inefficacy in Understanding Converse Relations</i>	
Chengwen Qi, Bowen Li, Binyuan Hui, Bailin Wang, Jinyang Li, Jinwang Wu and Yuanjun Laili	6932
<i>Towards Low-Resource Automatic Program Repair with Meta-Learning and Pretrained Language Models</i>	
Weishi Wang, Yue Wang, Steven Hoi and Shafiq Joty	6954
<i>ZGUL: Zero-shot Generalization to Unseen Languages using Multi-source Ensembling of Language Adapters</i>	
Vipul Kumar Rathore, Rajdeep Dhingra, Parag Singla and Mausam	6969
<i>Log-FGAER: Logic-Guided Fine-Grained Address Entity Recognition from Multi-Turn Spoken Dialogue</i>	
Xue Han, Yitong Wang, Qian Hu, Pengwei Hu, Chao Deng and Junlan Feng	6988
<i>Unified Low-Resource Sequence Labeling by Sample-Aware Dynamic Sparse Finetuning</i>	
Sarkar Snigdha Sarathi Das, Haoran Ranran Zhang, Peng Shi, Wenpeng Yin and Rui Zhang	6998
<i>On the Representational Capacity of Recurrent Neural Language Models</i>	
Franz Nowak, Anej Svete, Li Du and Ryan Cotterell	7011
<i>A Mechanistic Interpretation of Arithmetic Reasoning in Language Models using Causal Mediation Analysis</i>	
Alessandro Stolfo, Yonatan Belinkov and Mrinmaya Sachan	7035
<i>Benchmarking and Improving Text-to-SQL Generation under Ambiguity</i>	
Adithya Bhaskar, Tushar Tomar, Ashutosh Sathe and Sunita Sarawagi	7053
<i>Non-autoregressive Text Editing with Copy-aware Latent Alignments</i>	
Yu Zhang, Yue Zhang, Leyang Cui and Guohong Fu	7075
<i>Translating away Translationese without Parallel Data</i>	
Rricha Jalota, Koel Dutta Chowdhury, Cristina España-Bonet and Josef van Genabith	7086
<i>Prompt-Based Monte-Carlo Tree Search for Goal-oriented Dialogue Policy Planning</i>	
Xiao Yu, Maximillian Chen and Zhou Yu	7101
<i>UniMath: A Foundational and Multimodal Mathematical Reasoner</i>	
Zhenwen Liang, Tianyu Yang, Jipeng Zhang and Xiangliang Zhang	7126
<i>CaseEncoder: A Knowledge-enhanced Pre-trained Model for Legal Case Encoding</i>	
Yixiao Ma, Yueyue WU, Weihang Su, Qingyao Ai and Yiqun Liu	7134
<i>HiddenTables and PyQTax: A Cooperative Game and Dataset For TableQA to Ensure Scale and Data Privacy Across a Myriad of Taxonomies</i>	
William Watson, Nicole Cho, Tucker Balch and Manuela Veloso	7144
<i>Causal Document-Grounded Dialogue Pre-training</i>	
Yingxiu Zhao, Bowen Yu, Bowen Li, Haiyang Yu, Jinyang Li, Chao Wang, Fei Huang, Yongbin Li and Nevin L. Zhang	7160
<i>Accented Speech Recognition With Accent-specific Codebooks</i>	
Darshan Deepak Prabhu, Preethi Jyothi, Sriram Ganapathy and Vinit Unni	7175
<i>Linking Surface Facts to Large-Scale Knowledge Graphs</i>	
Gorjan Radevski, Kiril Gashteovski, Chia-Chien Hung, Carolin Lawrence and Goran Glavač	7189

<i>Sentiment Analysis on Streaming User Reviews via Dual-Channel Dynamic Graph Neural Network</i>	
Xin Zhang, Linhai Zhang and Deyu Zhou	7208
<i>DUMB: A Dutch Model Benchmark</i>	
Wietse de Vries, Martijn Wieling and Malvina Nissim	7221
<i>OssCSE: Overcoming Surface Structure Bias in Contrastive Learning for Unsupervised Sentence Embedding</i>	
Zhan Shi, Guoyin Wang, Ke Bai, Jiwei Li, Xiang Li, Qingjun Cui, Belinda Zeng, Trishul Chilimbi and Xiaodan Zhu	7242
<i>End-to-End Single-Channel Speaker-Turn Aware Conversational Speech Translation</i>	
Juan Pablo Zuluaga-Gomez, Zhaocheng Huang, Xing Niu, Rohit Paturi, Sundararajan Srinivasan, Prashant Mathur, Brian Thompson and Marcello Federico	7255
<i>A Fine-Grained Taxonomy of Replies to Hate Speech</i>	
Xinchen Yu, Ashley Zhao, Eduardo Blanco and Lingzi Hong	7275
<i>JointMatch: A Unified Approach for Diverse and Collaborative Pseudo-Labeling to Semi-Supervised Text Classification</i>	
Henry Peng Zou and Cornelia Caragea	7290
<i>Simple Temporal Adaptation to Changing Label Sets: Hashtag Prediction via Dense KNN</i>	
Niloofer Miresghallah, Nikolai Vogler, Junxian He, Omar Florez, Ahmed El-Kishky and Taylor Berg-Kirkpatrick	7302
<i>Speak, Memory: An Archaeology of Books Known to ChatGPT/GPT-4</i>	
Kent K. Chang, Mackenzie Hanh Cramer, Sandeep Soni and David Bamman	7312
<i>A Study on Accessing Linguistic Information in Pre-Trained Language Models by Using Prompts</i>	
Marion Di Marco, Katharina Hämmerl and Alexander Fraser	7328
<i>CiteBench: A Benchmark for Scientific Citation Text Generation</i>	
Martin Funkquist, Ilia Kuznetsov, Yufang Hou and Iryna Gurevych	7337
<i>From Heuristic to Analytic: Cognitively Motivated Strategies for Coherent Physical Commonsense Reasoning</i>	
Zheyuan Zhang, Shane Storks, Fengyuan Hu, Sungryull Sohn, Moontae Lee, Honglak Lee and Joyce Chai	7354
<i>A Challenging Multimodal Video Summary: Simultaneously Extracting and Generating Keyframe-Caption Pairs from Video</i>	
Keito Kudo, Haruki Nagasawa, Jun Suzuki and Nobuyuki Shimizu	7380
<i>Copyright Violations and Large Language Models</i>	
Antonia Karamolegkou, Jiaang Li, Li Zhou and Anders Søgaard	7403
<i>Effects of sub-word segmentation on performance of transformer language models</i>	
Jue Hou, Anisia Katinskaia, Anh-Duc Vu and Roman Yangarber	7413
<i>Symbolic Planning and Code Generation for Grounded Dialogue</i>	
Justin T Chiu, Wenting Zhao, Derek Chen, Saujas Vaduguru, Alexander M Rush and Daniel Fried	7426
<i>Universal Self-Adaptive Prompting</i>	
Xingchen Wan, Ruoxi Sun, Hootan Nakhost, Hanjun Dai, Julian Martin Eisenschlos, Sercan O Arik and Tomas Pfister	7437

<i>Somali Information Retrieval Corpus: Bridging the Gap between Query Translation and Dedicated Language Resources</i>	
Abdisalam Mahamed Badel, Ting Zhong, Wenxin Tai and Fan Zhou	7463
<i>Beat LLMs at Their Own Game: Zero-Shot LLM-Generated Text Detection via Querying ChatGPT</i>	
Biru Zhu, Lifan Yuan, Ganqu Cui, Yangyi Chen, Chong Fu, Bingxiang He, Yangdong Deng, Zhiyuan Liu, Maosong Sun and Ming Gu	7470
<i>Faithful Model Evaluation for Model-Based Metrics</i>	
Qian Hu, Palash Goyal and Rahul Gupta	7484
<i>Content- and Topology-Aware Representation Learning for Scientific Multi-Literature</i>	
Kai Zhang, Kaisong Song, Yangyang Kang and Xiaozhong Liu	7490
<i>Language Model Quality Correlates with Psychometric Predictive Power in Multiple Languages</i>	
Ethan Wilcox, Clara Meister, Ryan Cotterell and Tiago Pimentel	7503
<i>Joint Entity and Relation Extraction with Span Pruning and Hypergraph Neural Networks</i>	
Zhaohui Yan, Songlin Yang, Wei Liu and Kewei Tu	7512
<i>Have LLMs Advanced Enough? A Challenging Problem Solving Benchmark For Large Language Models</i>	
Daman Arora, Himanshu Gaurav Singh and Mausam	7527
<i>StrAE: Autoencoding for Pre-Trained Embeddings using Explicit Structure</i>	
Mattia Oppel, Victor Prokhorov and Siddharth N.	7544
<i>WiCE: Real-World Entailment for Claims in Wikipedia</i>	
Ryo Kamoi, Tanya Goyal, Juan Diego Rodriguez and Greg Durrett	7561
<i>Natural Disaster Tweets Classification Using Multimodal Data</i>	
Mohammad Abdul Basit, Bashir Alam, Zubaida Fatima and Salman Shaikh	7584
<i>On the Challenges of Using Black-Box APIs for Toxicity Evaluation in Research</i>	
Luiza Amador Pozzobon, Beyza Ermis, Patrick Lewis and Sara Hooker	7595
<i>RoBoCoP: A Comprehensive ROmance BORrowing COgnate Package and Benchmark for Multilingual Cognate Identification</i>	
Liviu P Dinu, Ana Sabina Uban, Alina Maria Cristea, Anca Daniela Dinu, Ioan-Bogdan Iordache, Simona Georgescu and Laurentiu Zoicas	7610
<i>Instructive Dialogue Summarization with Query Aggregations</i>	
Bin Wang, Zhengyuan Liu and Nancy F. Chen	7630
<i>Semantic matching for text classification with complex class descriptions</i>	
Brian M De Silva, Kuan-Wen Huang, Gwang Gook Lee, Karen Hovsepian, Yan Xu and Mingwei Shen	7654
<i>MADNet: Maximizing Addressee Deduction Expectation for Multi-Party Conversation Generation</i>	
Jia-Chen Gu, Chao-Hong Tan, Caiyuan Chu, Zhen-Hua Ling, Chongyang Tao, Quan Liu and Cong Liu	7681
<i>GLEN: Generative Retrieval via Lexical Index Learning</i>	
Sunkyoung Lee, Minjin Choi and Jongwuk Lee	7693
<i>Turn-Level Active Learning for Dialogue State Tracking</i>	
Zihan Zhang, Meng Fang, Fanghua Ye, Ling Chen and Mohammad-Reza Namazi-Rad	7705

<i>ReSee: Responding through Seeing Fine-grained Visual Knowledge in Open-domain Dialogue</i>	
Haoqin Tu, Yitong Li, Fei Mi and Zhongliang Yang	7720
<i>Modeling Conceptual Attribute Likeness and Domain Inconsistency for Metaphor Detection</i>	
Yuan Tian, Nan Xu, Wenji Mao and Daniel Dajun Zeng	7736
<i>Referring Image Segmentation via Joint Mask Contextual Embedding Learning and Progressive Alignment Network</i>	
Ziling Huang and Shin'ichi Satoh	7753
<i>Shall We Pretrain Autoregressive Language Models with Retrieval? A Comprehensive Study</i>	
Boxin Wang, Wei Ping, Peng Xu, Lawrence McAfee, Zihan Liu, Mohammad Shoeybi, Yi Dong, Oleksii Kuchaiev, Bo Li, Chaowei Xiao, Anima Anandkumar and Bryan Catanzaro	7763
<i>SCITAB: A Challenging Benchmark for Compositional Reasoning and Claim Verification on Scientific Tables</i>	
Xinyuan Lu, Liangming Pan, Qian Liu, Preslav Nakov and Min-Yen Kan	7787
<i>Training Simultaneous Speech Translation with Robust and Random Wait-k-Tokens Strategy</i>	
Linlin Zhang, Kai Fan, Jiajun Bu and Zhongqiang Huang	7814
<i>SCENE: Self-Labeled Counterfactuals for Extrapolating to Negative Examples</i>	
Deqing Fu, Ameya Godbole and Robin Jia	7832
<i>Enhancing Code-Switching for Cross-lingual SLU: A Unified View of Semantic and Grammatical Coherence</i>	
Zhihong Zhu, Xuxin Cheng, Zhiqi Huang, Dongsheng Chen and Yuexian Zou	7849
<i>Task-Agnostic Low-Rank Adapters for Unseen English Dialects</i>	
Zedian Xiao, William Barr Held, Yanchen Liu and Diyi Yang	7857
<i>Federated Learning of Large Language Models with Parameter-Efficient Prompt Tuning and Adaptive Optimization</i>	
Tianshi Che, Ji Liu, Yang Zhou, Jiaxiang Ren, Jiwen Zhou, Victor S. Sheng, Huaiyu Dai and Dejing Dou	7871
<i>TheoremQA: A Theorem-driven Question Answering Dataset</i>	
Wenhu Chen, Ming Yin, Max Ku, Pan Lu, Yixin Wan, Xueguang Ma, Jianyu Xu, Xinyi Wang and Tony Xia	7889
<i>Scalable-DSC: A Structural Template Prompt Approach to Scalable Dialogue State Correction</i>	
Haoxiang Su, Hongyan Xie, Hao Huang, Shuangyong Song, Ruiyu Fang, Xiaomeng Huang and Sijie Feng	7902
<i>Don't Trust ChatGPT when your Question is not in English: A Study of Multilingual Abilities and Types of LLMs</i>	
Xiang Zhang, Senyu Li, Bradley Hauer, Ning Shi and Grzegorz Kondrak	7915
<i>M³Seg: A Maximum-Minimum Mutual Information Paradigm for Unsupervised Topic Segmentation in ASR Transcripts</i>	
Ke Wang, Xiutian Zhao, Yanghui Li and Wei Peng	7928
<i>Empirical Study of Zero-Shot NER with ChatGPT</i>	
Tingyu Xie, Qi Li, Jian Zhang, Yan Zhang, Zuozhu Liu and Hongwei Wang	7935
<i>Automatic Prompt Optimization with "Gradient Descent" and Beam Search</i>	
Reid Pryzant, Dan Iter, Jerry Li, Yin Tat Lee, Chenguang Zhu and Michael Zeng	7957

<i>Active Retrieval Augmented Generation</i>	
Zhengbao Jiang, Frank F. Xu, Luyu Gao, Zhiqing Sun, Qian Liu, Jane Dwivedi-Yu, Yiming Yang, Jamie Callan and Graham Neubig	7969
<i>GD-COMET: A Geo-Diverse Commonsense Inference Model</i>	
Mehar Bhatia and Vered Shwartz	7993
<i>Multi-level Adaptive Contrastive Learning for Knowledge Internalization in Dialogue Generation</i>	
Chenxu Yang, Zheng Lin, Lanrui Wang, Chong Tian, Liang Pang, Jiangnan Li, Qirong Ho, Yanan Cao and Weiping Wang	8002
<i>Enhancing Biomedical Lay Summarisation with External Knowledge Graphs</i>	
Tomas Goldsack, Zhihao Zhang, Chen Tang, Carolina Scarton and Chenghua Lin	8016
<i>A Diffusion Weighted Graph Framework for New Intent Discovery</i>	
Wenkai Shi, Wenbin An, Feng Tian, Qinghua Zheng, QianYing Wang and Ping Chen	8033
<i>A Self-enhancement Multitask Framework for Unsupervised Aspect Category Detection</i>	
Thi-Nhung Nguyen, Hoang Ngo, Kiem-Hieu Nguyen and Tuan-Dung Cao	8043
<i>DialCoT Meets PPO: Decomposing and Exploring Reasoning Paths in Smaller Language Models</i>	
Chengcheng Han, Xiaowei Du, Che Zhang, Yixin Lian, Xiang Li, Ming Gao and Baoyuan Wang	8055
<i>Recurrent Neural Language Models as Probabilistic Finite-state Automata</i>	
Anej Svete and Ryan Cotterell	8069
<i>Revisiting Source Context in Nearest Neighbor Machine Translation</i>	
Xuanhong Li, Peng Li and Po Hu	8087
<i>Find-2-Find: Multitask Learning for Anaphora Resolution and Object Localization</i>	
Cennet Oguz, Pascal Denis, Emmanuel Vincent, Simon Ostermann and Josef van Genabith	8099
<i>Background Summarization of Event Timelines</i>	
Adithya Pratapa, Kevin Small and Markus Dreyer	8111
<i>Superlim: A Swedish Language Understanding Evaluation Benchmark</i>	
Aleksandrs Berdicevskis, Gerlof Bouma, Robin Kurtz, Felix Morger, Joey Öhman, Yvonne Adesam, Lars Borin, Dana Dannélls, Markus Forsberg, Tim Isbister, Anna Lindahl, Martin Malmsten, Faton Rekathati, Magnus Sahlgren, Elena Volodina, Love Börjeson, Simon Hengchen and Nina Tahmasebi	8137
<i>Reasoning with Language Model is Planning with World Model</i>	
Shibo Hao, Yi Gu, Haodi Ma, Joshua Jiahua Hong, Zhen Wang, Daisy Zhe Wang and Zhiting Hu	8154
<i>LLM-enhanced Self-training for Cross-domain Constituency Parsing</i>	
Jianling Li, Meishan Zhang, Peiming Guo, Min Zhang and Yue Zhang	8174
<i>Continual Named Entity Recognition without Catastrophic Forgetting</i>	
Duzhen Zhang, Wei Cong, Jiahua Dong, Yahan Yu, Xiuyi Chen, Yonggang Zhang and Zhen Fang	8186
<i>DSI++: Updating Transformer Memory with New Documents</i>	
Sanket Vaibhav Mehta, Jai Gupta, Yi Tay, Mostafa Dehghani, Vinh Q. Tran, Jinfeng Rao, Marc Najork, Emma Strubell and Donald Metzler	8198

<i>Editing Common Sense in Transformers</i>	
Anshita Gupta, Debanjan Mondal, Akshay Krishna Sheshadri, Wenlong Zhao, Xiang Lorraine Li, Sarah Wiegrefe and Niket Tandon	8214
<i>Air-Decoding: Attribute Distribution Reconstruction for Decoding-Time Controllable Text Generation</i>	
Tianqi Zhong, Quan Wang, Jingxuan Han, Yongdong Zhang and Zhendong Mao	8233
<i>Homophone Disambiguation Reveals Patterns of Context Mixing in Speech Transformers</i>	
Hosein Mohebbi, Grzegorz Chrupala, Willem Zuidema and Afra Alishahi	8249
<i>Retrieval-Generation Alignment for End-to-End Task-Oriented Dialogue System</i>	
Weizhou Shen, Yingqi Gao, Canbin Huang, Fanqi Wan, Xiaojun Quan and Wei Bi	8261
<i>IfQA: A Dataset for Open-domain Question Answering under Counterfactual Presuppositions</i>	
Wenhao Yu, Meng Jiang, Peter Clark and Ashish Sabharwal	8276
<i>How Do Large Language Models Capture the Ever-changing World Knowledge? A Review of Recent Advances</i>	
Zihan Zhang, Meng Fang, Ling Chen, Mohammad-Reza Namazi-Rad and Jun Wang	8289
<i>PreWoMe: Exploiting Presuppositions as Working Memory for Long Form Question Answering</i>	
Wookje Han, Jinsol Park and Kyungjae Lee	8312
<i>Memorisation Cartography: Mapping out the Memorisation-Generalisation Continuum in Neural Machine Translation</i>	
Verna Dankers, Ivan Titov and Dieuwke Hupkes	8323
<i>DecipherPref: Analyzing Influential Factors in Human Preference Judgments via GPT-4</i>	
Yebowen Hu, Kaiqiang Song, Sangwoo Cho, Xiaoyang Wang, Hassan Foroosh and Fei Liu	8344
<i>Gender Biases in Automatic Evaluation Metrics for Image Captioning</i>	
Haoyi Qiu, Zi-Yi Dou, Tianlu Wang, Asli Celikyilmaz and Nanyun Peng	8358
<i>QA-NatVer: Question Answering for Natural Logic-based Fact Verification</i>	
Rami Aly, Marek Strong and Andreas Vlachos	8376
<i>Increasing Probability Mass on Answer Choices Does Not Always Improve Accuracy</i>	
Sarah Wiegrefe, Matthew Finlayson, Oyvind Tafjord, Peter Clark and Ashish Sabharwal	8392
<i>Generating Data for Symbolic Language with Large Language Models</i>	
Jiacheng Ye, Chengzu Li, Lingpeng Kong and Tao Yu	8418
<i>IDTraffickers: An Authorship Attribution Dataset to link and connect Potential Human-Trafficking Operations on Text Escort Advertisements</i>	
Vageesh Kumar Saxena, Benjamin Ashpole, Gijs van Dijck and Gerasimos Spanakis	8444
<i>Evaluating Bias and Fairness in Gender-Neutral Pretrained Vision-and-Language Models</i>	
Laura Cabello, Emanuele Bugliarello, Stephanie Brandl and Desmond Elliott	8465
<i>Improving Dialogue Discourse Parsing via Reply-to Structures of Addressee Recognition</i>	
Yaxin Fan, Feng Jiang, Peifeng Li, Fang Kong and Qiaoming Zhu	8484
<i>Improving Language Models' Meaning Understanding and Consistency by Learning Conceptual Roles from Dictionary</i>	
Myeongjun Erik Jang and Thomas Lukasiewicz	8496

<i>DALE: Generative Data Augmentation for Low-Resource Legal NLP</i>	
Sreyan Ghosh, Chandra Kiran Reddy Evuru, Sonal Kumar, S Ramaneswaran, S Sakshi, Utkarsh Tyagi and Dinesh Manocha	8511
<i>FedID: Federated Interactive Distillation for Large-Scale Pretraining Language Models</i>	
Xinge Ma, Jiangming Liu, Jin Wang and Xuejie Zhang	8566
<i>trlX: A Framework for Large Scale Reinforcement Learning from Human Feedback</i>	
Alexander Havrilla, Maksym Zhuravinskiy, Duy Van Phung, Aman Tiwari, Jonathan Tow, Stella Biderman, Quentin Gregory Anthony and Louis Castricato	8578
<i>This is not a Dataset: A Large Negation Benchmark to Challenge Large Language Models</i>	
Iker García-Ferrero, Begoña Altuna, Javier Alvez, Itziar Gonzalez-Dios and German Rigau	8596
<i>MT2: Towards a Multi-Task Machine Translation Model with Translation-Specific In-Context Learning</i>	
Chunyou Li, Mingtong Liu, Hongxiao Zhang, Yufeng Chen, Jinan Xu and Ming Zhou	8616
<i>CleanCoNLL: A Nearly Noise-Free Named Entity Recognition Dataset</i>	
Susanna Rücker and Alan Akbik	8628
<i>Disentangling Transformer Language Models as Superposed Topic Models</i>	
Jia Peng Lim and Hady W. Lauw	8646
<i>Conversational Semantic Parsing using Dynamic Context Graphs</i>	
Parag Jain and Mirella Lapata	8667
<i>Not all quantifiers are equal: Probing Transformer-based language models' understanding of generalised quantifiers</i>	
Tharindu Madusanka, Iqra Zahid, Hao Li, Ian Pratt-Hartmann and Riza Batista-Navarro ...	8680
<i>Structure-aware Knowledge Graph-to-text Generation with Planning Selection and Similarity Distinction</i>	
Feng Zhao, Hongzhi Zou and Cheng Yan	8693
<i>SOUL: Towards Sentiment and Opinion Understanding of Language</i>	
Yue Deng, Wenxuan Zhang, Sinno Jialin Pan and Lidong Bing	8704
<i>Regulation and NLP (RegNLP): Taming Large Language Models</i>	
Catalina Goanta, Nikolaos Aletras, Ilias Chalkidis, Sofia Ranchordás and Gerasimos Spanakis	8712
<i>MedEval: A Multi-Level, Multi-Task, and Multi-Domain Medical Benchmark for Language Model Evaluation</i>	
Zexue He, Yu Wang, An Yan, Yao Liu, Eric Y Chang, Amilcare Gentili, Julian McAuley and Chun-Nan Hsu	8725
<i>Seeing through the mess: evolutionary dynamics of lexical polysemy</i>	
Andreas Baumann, Andreas Stephan and Benjamin Roth	8745
<i>Are Embedded Potatoes Still Vegetables? On the Limitations of WordNet Embeddings for Lexical Semantics</i>	
Xuyou Cheng, Michael Sejr Schlichtkrull and Guy Emerson	8763
<i>Evaluation Metrics in the Era of GPT-4: Reliably Evaluating Large Language Models on Sequence to Sequence Tasks</i>	
Andrea Sottana, Bin Liang, Kai Zou and Zheng Yuan	8776

<i>Event-Location Tracking in Narratives: A Case Study on Holocaust Testimonies</i> Eitan Wagner, Renana Keydar and Omri Abend	8789
<i>Dialogizer: Context-aware Conversational-QA Dataset Generation from Textual Sources</i> Yerin Hwang, Yongil Kim, Hyunkyung Bae, Hwanhee Lee, Jeesoo Bang and Kyomin Jung ..	8806
<i>Learning to Predict Task Transferability via Soft Prompt</i> Lingyun Feng	8829
<i>Chain-of-Questions Training with Latent Answers for Robust Multistep Question Answering</i> Wang Zhu, Jesse Thomason and Robin Jia	8845
<i>Mirror: A Universal Framework for Various Information Extraction Tasks</i> Tong Zhu, Junfei Ren, Zijian Yu, Mengsong Wu, Guoliang Zhang, Xiaoye Qu, Wenliang Chen, Zhefeng Wang, Baoxing Huai and Min Zhang	8861
<i>“Mistakes Help Us Grow”: Facilitating and Evaluating Growth Mindset Supportive Language in Classrooms</i> Kunal Handa, Margaret Clapper, Jessica Boyle, Rose E Wang, Diyi Yang, David Yeager and Dorottya Demszky	8877
<i>Unnatural Error Correction: GPT-4 Can Almost Perfectly Handle Unnatural Scrambled Text</i> Qi Cao, Takeshi Kojima, Yutaka Matsuo and Yusuke Iwasawa	8898
<i>Detecting and Mitigating Hallucinations in Multilingual Summarisation</i> Yifu Qiu, Yftah Ziser, Anna Korhonen, Edoardo Ponti and Shay B. Cohen	8914
<i>Exploring Linguistic Probes for Morphological Inflection</i> Jordan Kodner, Salam Khalifa and Sarah Ruth Brogden Payne	8933
<i>AMR Parsing with Causal Hierarchical Attention and Pointers</i> Chao Lou and Kewei Tu	8942
<i>FLatS: Principled Out-of-Distribution Detection with Feature-Based Likelihood Ratio Score</i> Haowei Lin and Yuntian Gu	8956
<i>Self-Evolution Learning for Mixup: Enhance Data Augmentation on Few-Shot Text Classification Tasks</i> Haoqi Zheng, Qihuang Zhong, Liang Ding, Zhiliang Tian, Xin Niu, Changjian Wang, Dongsheng Li and Dacheng Tao	8964
<i>IC3: Image Captioning by Committee Consensus</i> David Chan, Austin Myers, Sudheendra Vijayanarasimhan, David A Ross and John Canny ..	8975
<i>SelfCheckGPT: Zero-Resource Black-Box Hallucination Detection for Generative Large Language Models</i> Potsawee Manakul, Adian Liusie and Mark Gales	9004
<i>Fair Without Leveling Down: A New Intersectional Fairness Definition</i> Gaurav Maheshwari, Aurélien Bellet, Pascal Denis and Mikaela Keller	9018
<i>Revisiting Instruction Fine-tuned Model Evaluation to Guide Industrial Applications</i> Manuel Faysse, Gautier Viaud, Céline Hudelot and Pierre Colombo	9033
<i>CLAD-ST: Contrastive Learning with Adversarial Data for Robust Speech Translation</i> Sathish Reddy Indurthi, Shamil Chollampatt, Ravi Agrawal and Marco Turchi	9049

<i>M2DF: Multi-grained Multi-curriculum Denoising Framework for Multimodal Aspect-based Sentiment Analysis</i>	
Fei Zhao, Chunhui Li, Zhen Wu, Yawen Ouyang, Jianbing Zhang and Xinyu Dai	9057
<i>Detection of Multiple Mental Disorders from Social Media with Two-Stream Psychiatric Experts</i>	
Si yuan Chen, Zhiling Zhang, Mengyue Wu and Kenny Q. Zhu	9071
<i>Understanding the Role of Input Token Characters in Language Models: How Does Information Loss Affect Performance?</i>	
Ahmed Alajrami, Katerina Margatina and Nikolaos Aletras	9085
<i>Improved Unsupervised Chinese Word Segmentation Using Pre-trained Knowledge and Pseudo-labeling Transfer</i>	
Hsiu-Wen Li, Ying-Jia Lin, Yi-Ting Li, Chun Yi Lin and Hung-Yu Kao	9109
<i>EasyQuant: An Efficient Data-free Quantization Algorithm for LLMs</i>	
Hanlin Tang, Yifu Sun, Decheng Wu, Kai Liu, Jianchen Zhu and Zhanhui Kang	9119
<i>Polar Ducks and Where to Find Them: Enhancing Entity Linking with Duck Typing and Polar Box Embeddings</i>	
Mattia Atzeni, Mikhail Plekhanov, Frederic A Dreyer, Nora Kassner, Simone Merello, Louis Martin and Nicola Cancedda	9129
<i>APrompt: Attention Prompt Tuning for Efficient Adaptation of Pre-trained Language Models</i>	
Qifan Wang, Yuning Mao, Jingang Wang, Hanchao Yu, Shaoliang Nie, Sinong Wang, Fuli Feng, Lifu Huang, Xiaojun Quan, Zenglin Xu and Dongfang Liu	9147
<i>What’s “up” with vision-language models? Investigating their struggle with spatial reasoning</i>	
Amita Kamath, Jack Hessel and Kai-Wei Chang	9161
<i>IBADR: an Iterative Bias-Aware Dataset Refinement Framework for Debiasing NLU models</i>	
Xiaoyue Wang, Xin Liu, Lijie Wang, Yaoxiang Wang, Jinsong Su and Hua Wu	9176
<i>Learning Preference Model for LLMs via Automatic Preference Data Generation</i>	
Shijia Huang, Jianqiao Zhao, Yanyang Li and Liwei Wang	9187
<i>Multilingual k-Nearest-Neighbor Machine Translation</i>	
David Stap and Christof Monz	9200
<i>Understanding Computational Models of Semantic Change: New Insights from the Speech Community</i>	
Filip Miletic, Anne Przewozny-Desriaux and Ludovic Tanguy	9209
<i>Causal Reasoning through Two Cognition Layers for Improving Generalization in Visual Question Answering</i>	
Trang Nguyen and Naoaki Okazaki	9221
<i>StructGPT: A General Framework for Large Language Model to Reason over Structured Data</i>	
Jinhao Jiang, Kun Zhou, Zican Dong, Keming Ye, Xin Zhao and Ji-Rong Wen	9237
<i>Modeling Legal Reasoning: LM Annotation at the Edge of Human Agreement</i>	
Rosamond Elizabeth Thalken, Edward Stiglitz, David Mimno and Matthew Wilkens	9252
<i>Model-tuning Via Prompts Makes NLP Models Adversarially Robust</i>	
Mrigank Raman, Pratyush Maini, J Zico Kolter, Zachary Chase Lipton and Danish Pruthi	9266
<i>Learning Co-Speech Gesture for Multimodal Aphasia Type Detection</i>	
Daeun Lee, Sejung Son, Hyolim Jeon, Seungbae Kim and Jinyoung Han	9287

<i>STINMatch: Semi-Supervised Semantic-Topological Iteration Network for Financial Risk Detection via News Label Diffusion</i>	
Xurui Li, Yue Qin, Rui Zhu, Tianqianjin Lin, Yongming Fan, Yangyang Kang, Kaisong Song, Fubang Zhao, Changlong Sun, Haixu Tang and Xiaozhong Liu	9304
<i>Centering the Margins: Outlier-Based Identification of Harmed Populations in Toxicity Detection</i>	
Vyoma Raman, Eve Fleisig and Dan Klein	9316
<i>Describe Me an Auklet: Generating Grounded Perceptual Category Descriptions</i>	
Bill Noble and Nikolai Ilinykh	9330
<i>Revisiting Automated Topic Model Evaluation with Large Language Models</i>	
Dominik Stambach, Vilém Zouhar, Alexander Hoyle, Mrinmaya Sachan and Elliott Ash . .	9348
<i>ORCHID: A Chinese Debate Corpus for Target-Independent Stance Detection and Argumentative Dialogue Summarization</i>	
Xiutian Zhao, Ke Wang and Wei Peng	9358
<i>On the Benefits of Learning to Route in Mixture-of-Experts Models</i>	
Nishanth Dikkala, Nikhil Ghosh, Raghu Meka, Rina Panigrahy, Nikhil Vyas and Xin Wang	9376
<i>SEAHORSE: A Multilingual, Multifaceted Dataset for Summarization Evaluation</i>	
Elizabeth Clark, Shruti Rijhwani, Sebastian Gehrmann, Joshua Maynez, Roei Aharoni, Vitaly Nikolaev, Thibault Sellam, Aditya Siddhant, Dipanjan Das and Ankur P Parikh	9397
<i>Query2doc: Query Expansion with Large Language Models</i>	
Liang Wang, Nan Yang and Furu Wei	9414
<i>We Need to Talk About Reproducibility in NLP Model Comparison</i>	
Yan Xue, Xuefei Cao, Xingli Yang, Yu Wang, Ruibo Wang and Jihong Li	9424
<i>Explore-Instruct: Enhancing Domain-Specific Instruction Coverage through Active Exploration</i>	
Fanqi Wan, Xinting Huang, Tao Yang, Xiaojun Quan, Wei Bi and Shuming Shi	9435
<i>Practical Computational Power of Linear Transformers and Their Recurrent and Self-Referential Extensions</i>	
Kazuki Irie, Róbert Csordás and Jürgen Schmidhuber	9455
<i>InterFair: Debiasing with Natural Language Feedback for Fair Interpretable Predictions</i>	
Bodhisattwa Prasad Majumder, Zexue He and Julian McAuley	9466
<i>Just Adjust One Prompt: Enhancing In-Context Dialogue Scoring via Constructing the Optimal Sub-graph of Demonstrations and Prompts</i>	
Jiashu Pu, Ling Cheng, Lu Fan, Tangjie Lv and Rongsheng Zhang	9472
<i>Multilingual estimation of political-party positioning: From label aggregation to long-input Transformers</i>	
Dmitry Nikolaev, Tanise Ceron and Sebastian Padó	9497
<i>ART: rule bAsed futuRe-inference deducTion</i>	
Mengze Li, Tianqi Zhao, Bai Jionghao, Baoyi He, Jiaxu Miao, Wei Ji, Zheqi Lv, Zhou Zhao, Shengyu Zhang, Wenqiao Zhang and Fei Wu	9512
<i>EpiK-Eval: Evaluation for Language Models as Epistemic Models</i>	
Gabriele Prato, Jerry Huang, Prasanna Parthasarathi, Shagun Sodhani and Sarath Chandar .	9523

<i>From Dissonance to Insights: Dissecting Disagreements in Rationale Construction for Case Outcome Classification</i>	
Shanshan Xu, Santosh T.Y.S.S, Oana Ichim, Isabella Risini, Barbara Plank and Matthias Grabmair	9558
<i>On Bilingual Lexicon Induction with Large Language Models</i>	
Yaoyiran Li, Anna Korhonen and Ivan Vulic	9577
<i>Statistical Depth for Ranking and Characterizing Transformer-Based Text Embeddings</i>	
Parker Seegmiller and Sarah Masud Preum	9600
<i>CRaSh: Clustering, Removing, and Sharing Enhance Fine-tuning without Full Large Language Model</i>	
Kaiyan Zhang, Ning Ding, Biqing Qi, Xuekai Zhu, Xinwei Long and Bowen Zhou	9612
<i>From Multilingual Complexity to Emotional Clarity: Leveraging Commonsense to Unveil Emotions in Code-Mixed Dialogues</i>	
Shivani Kumar, Ramaneswaran S, Md Shad Akhtar and Tanmoy Chakraborty	9638
<i>Large Language Models are biased to overestimate profoundness</i>	
Eugenio Herrera-Berg, Tomás Vergara Browne, Pablo León-Villagrà, Marc-Lluís Vives and Cristian Buc Calderon	9653
<i>SummEdits: Measuring LLM Ability at Factual Reasoning Through The Lens of Summarization</i>	
Philippe Laban, Wojciech Maciej Kryscinski, Divyansh Agarwal, Alexander Fabbri, Caiming Xiong, Shafiq Joty and Chien-Sheng Wu	9662
<i>DIVE: Towards Descriptive and Diverse Visual Commonsense Generation</i>	
Jun-Hyung Park, Hyuntae Park, Youjin Kang, Eojin Jeon and SangKeun Lee	9677
<i>Towards Conceptualization of “Fair Explanation”: Disparate Impacts of anti-Asian Hate Speech Explanations on Content Moderators</i>	
Tin Trung Nguyen, Jiannan Xu, Aayushi Roy, Hal Daumé III and Marine Carpuat	9696
<i>Bridging Background Knowledge Gaps in Translation with Automatic Explication</i>	
HyoJung Han, Jordan Lee Boyd-Graber and Marine Carpuat	9718
<i>A Quality-based Syntactic Template Retriever for Syntactically-Controlled Paraphrase Generation</i>	
Xue Zhang, Songming Zhang, Yunlong Liang, Yufeng Chen, Jian Liu, Wenjuan Han and Jinan Xu	9736
<i>Beyond Shared Vocabulary: Increasing Representational Word Similarities across Languages for Multilingual Machine Translation</i>	
Di Wu and Christof Monz	9749
<i>Quantifying the redundancy between prosody and text</i>	
Lukas Wolf, Tiago Pimentel, Evelina Fedorenko, Ryan Cotterell, Alex Warstadt, Ethan Wilcox and Tamar I Regev	9765
<i>CRoW: Benchmarking Commonsense Reasoning in Real-World Tasks</i>	
Mete Ismayilzada, Debjit Paul, Syrielle Montariol, Mor Geva and Antoine Bosselut	9785
<i>A Video Is Worth 4096 Tokens: Verbalize Story Videos To Understand Them In Zero Shot</i>	
Aanisha Bhattacharyya, Yaman K Singla, Balaji Krishnamurthy, Rajiv Ratn Shah and Changyou Chen	9822
<i>Label Words are Anchors: An Information Flow Perspective for Understanding In-Context Learning</i>	
Lean Wang, Lei Li, Damai Dai, Deli Chen, Hao Zhou, Fandong Meng, Jie Zhou and Xu Sun	9840

<i>Prompting Scientific Names for Zero-Shot Species Recognition</i>	
Shubham Parashar, Zhiqiu Lin, Yanan Li and Shu Kong	9856
<i>Active Learning for Natural Language Generation</i>	
Yotam Perlitz, Ariel Gera, Michal Shmueli-Scheuer, Dafna Sheinwald, Noam Slonim and Liat Ein-Dor	9862
<i>Re³Dial: Retrieve, Reorganize and Rescale Conversations for Long-Turn Open-Domain Dialogue Pre-training</i>	
Jiaxin Wen, Hao Zhou, Jian Guan, Jie Zhou and Minlie Huang	9878
<i>MultiTurnCleanup: A Benchmark for Multi-Turn Spoken Conversational Transcript Cleanup</i>	
Hua Shen, Vicky Zayats, Johann C Rocholl, Daniel David Walker and Dirk Padfield	9895
<i>Do All Languages Cost the Same? Tokenization in the Era of Commercial Language Models</i>	
Orevaoghene Ahia, Sachin Kumar, Hila Gonen, Jungo Kasai, David R Mortensen, Noah A. Smith and Yulia Tsvetkov	9904
<i>Characterizing Mechanisms for Factual Recall in Language Models</i>	
Qinan Yu, Jack Merullo and Ellie Pavlick	9924
<i>MULTITuDE: Large-Scale Multilingual Machine-Generated Text Detection Benchmark</i>	
Dominik Macko, Robert Moro, Adaku Uchendu, Jason S Lucas, Michiharu Yamashita, Matúš Pikuliak, Ivan Srba, Thai Le, Dongwon Lee, Jakub Simko and Maria Bielikova	9960
<i>Revisiting Block-based Quantisation: What is Important for Sub-8-bit LLM Inference?</i>	
Cheng Zhang, Jianyi Cheng, Ilia Shumailov, George Anthony Constantinides and Yiren Zhao	9988
<i>Whispering LLaMA: A Cross-Modal Generative Error Correction Framework for Speech Recognition</i>	
Srijith Radhakrishnan, Chao-Han Huck Yang, Sumeer Ahmad Khan, Rohit Kumar, Narsis A. Kiani, David Gomez-Cabrero and Jesper Tegnér	10007
<i>Reducing Sequence Length by Predicting Edit Spans with Large Language Models</i>	
Masahiro Kaneko and Naoaki Okazaki	10017
<i>Instruct and Extract: Instruction Tuning for On-Demand Information Extraction</i>	
Yizhu Jiao, Ming Zhong, Sha Li, Ruining Zhao, Siru Ouyang, Heng Ji and Jiawei Han	10030
<i>Rethinking the Evaluation for Conversational Recommendation in the Era of Large Language Models</i>	
Xiaolei Wang, Xinyu Tang, Xin Zhao, Jingyuan Wang and Ji-Rong Wen	10052
<i>ReCEval: Evaluating Reasoning Chains via Correctness and Informativeness</i>	
Archiki Prasad, Swarnadeep Saha, Xiang Zhou and Mohit Bansal	10066
<i>Expand, Highlight, Generate: RL-driven Document Generation for Passage Reranking</i>	
Arian Askari, Mohammad Aliannejadi, Chuan Meng, Evangelos Kanoulas and Suzan Verberne	10087
<i>Transformer-based Live Update Generation for Soccer Matches from Microblog Posts</i>	
Masashi Oshika, Kosuke Yamada, Ryohei Sasano and Koichi Takeda	10100
<i>Make Every Example Count: On the Stability and Utility of Self-Influence for Learning from Noisy NLP Datasets</i>	
Irina Bejan, Artem Sokolov and Katja Filippova	10107
<i>Appraising the Potential Uses and Harms of LLMs for Medical Systematic Reviews</i>	
Hye Sun Yun, Iain James Marshall, Thomas Trikalinos and Byron C Wallace	10122

<i>PromptST: Abstract Prompt Learning for End-to-End Speech Translation</i>	
Tengfei Yu, Liang Ding, Xuebo Liu, Kehai Chen, Meishan Zhang, Dacheng Tao and Min Zhang	
	10140
<i>Text Rendering Strategies for Pixel Language Models</i>	
Jonas F. Lotz, Elizabeth Salesky, Phillip Rust and Desmond Elliott	10155
<i>APoLLO : Unified Adapter and Prompt Learning for Vision Language Models</i>	
Sanjoy Chowdhury, Sayan Nag and Dinesh Manocha	10173
<i>SAMRank: Unsupervised Keyphrase Extraction using Self-Attention Map in BERT and GPT-2</i>	
Byungha Kang and Youhyun Shin	10188
<i>Contrastive Learning for Inference in Dialogue</i>	
Etsuko Ishii, Yan Xu, Bryan Wilie, Ziwei Ji, Holy Lovenia, Willy Chung and Pascale Fung	10202
<i>Editing Large Language Models: Problems, Methods, and Opportunities</i>	
Yunzhi Yao, Peng Wang, Bozhong Tian, Siyuan Cheng, Zhoubo Li, Shumin Deng, Huajun Chen and Ningyu Zhang	10222
<i>MarkQA: A large scale KBQA dataset with numerical reasoning</i>	
Xiang Huang, Sitao Cheng, Yuheng Bao, Shanshan Huang and Yuzhong Qu	10241
<i>Comparing Biases and the Impact of Multilingual Training across Multiple Languages</i>	
Sharon Levy, Neha Anna John, Ling Liu, Yogarshi Vyas, Jie Ma, Yoshinari Fujinuma, Miguel Ballesteros, Vittorio Castelli and Dan Roth	10260
<i>HutCRS: Hierarchical User-Interest Tracking for Conversational Recommender System</i>	
Mingjie Qian, Yongsen Zheng, Jinghui Qin and Liang Lin	10281
<i>Large Language Models Meet Open-World Intent Discovery and Recognition: An Evaluation of Chat-GPT</i>	
Xiaoshuai Song, Keqing He, Pei Wang, Guanting Dong, Yutao Mou, Jingang Wang, Yunsen Xian, Xunliang Cai and Weiran Xu	10291
<i>The Distributional Hypothesis Does Not Fully Explain the Benefits of Masked Language Model Pretraining</i>	
Ting-Rui Chiang and Dani Yogatama	10305
<i>Simple and Effective Input Reformulations for Translation</i>	
Brian Yu, Hansen Lillemark and Kurt Keutzer	10322
<i>Pointwise Mutual Information Based Metric and Decoding Strategy for Faithful Generation in Document Grounded Dialogs</i>	
Yatin Nandwani, Vineet Kumar, Dinesh Raghu, Sachindra Joshi and Luis A. Lastras	10335
<i>The ACL OCL Corpus: Advancing Open Science in Computational Linguistics</i>	
Shaurya Rohatgi, Yanxia Qin, Benjamin Aw, Niranjana Anand Unnithan and Min-Yen Kan	10348
<i>Using Artificial French Data to Understand the Emergence of Gender Bias in Transformer Language Models</i>	
Lina Conti and Guillaume Wisniewski	10362
<i>Learning to Rank Context for Named Entity Recognition Using a Synthetic Dataset</i>	
Arthur Amalvy, Vincent Labatut and Richard Dufour	10372

<i>Improving Diversity of Demographic Representation in Large Language Models via Collective-Critiques and Self-Voting</i>	
Preethi Lahoti, Nicholas Blumm, Xiao Ma, Raghavendra Kotikalapudi, Sahitya Potluri, Qijun Tan, Hansa Srinivasan, Ben Packer, Ahmad Beirami, Alex Beutel and Jilin Chen	10383
<i>Hidding the Ghostwriters: An Adversarial Evaluation of AI-Generated Student Essay Detection</i>	
Xinlin Peng, Ying Zhou, Ben He, Le Sun and Yingfei Sun	10406
<i>Contextual Interaction for Argument Post Quality Assessment</i>	
Yiran Wang, Xuanang Chen, Ben He and Le Sun	10420
<i>Pre-training Intent-Aware Encoders for Zero- and Few-Shot Intent Classification</i>	
Mujeen Sung, James Gung, Elman Mansimov, Nikolaos Pappas, Raphael Shu, Salvatore Romeo, Yi Zhang and Vittorio Castelli	10433
<i>Synthetic Data Generation with Large Language Models for Text Classification: Potential and Limitations</i>	
Zhuoyan Li, Hangxiao Zhu, Zhuoran Lu and Ming Yin	10443
<i>GazeVQA: A Video Question Answering Dataset for Multiview Eye-Gaze Task-Oriented Collaborations</i>	
Muhammet Furkan Ilaslan, Chenan Song, Joya Chen, Difei Gao, Weixian Lei, Qianli Xu, Joo Hwee Lim and Mike Zheng Shou	10462
<i>People Make Better Edits: Measuring the Efficacy of LLM-Generated Counterfactually Augmented Data for Harmful Language Detection</i>	
Indira Sen, Dennis Assenmacher, Mattia Samory, Isabelle Augenstein, Wil Aalst and Claudia Wagner	10480
<i>Unraveling Feature Extraction Mechanisms in Neural Networks</i>	
Xiaobing Sun, Jiaxi Li and Wei Lu	10505
<i>CAPSTONE: Curriculum Sampling for Dense Retrieval with Document Expansion</i>	
Xingwei He, Yeyun Gong, A-Long Jin, Hang Zhang, Anlei Dong, Jian Jiao, Siu Ming Yiu and Nan Duan	10531
<i>Balance Act: Mitigating Hubness in Cross-Modal Retrieval with Query and Gallery Banks</i>	
Yimu Wang, Xiangru Jian and Bo Xue	10542
<i>E-CORE: Emotion Correlation Enhanced Empathetic Dialogue Generation</i>	
Fengyi Fu, Lei Zhang, Quan Wang and Zhendong Mao	10568
<i>What do Deck Chairs and Sun Hats Have in Common? Uncovering Shared Properties in Large Concept Vocabularies</i>	
Amit Gajbhiye, Zied Bouraoui, Na Li, Usashi Chatterjee, Luis Espinosa-Anke and Steven Schoc-kaert	10587
<i>ALDi: Quantifying the Arabic Level of Dialectness of Text</i>	
Amr Keleg, Sharon Goldwater and Walid Magdy	10597
<i>3DRP-Net: 3D Relative Position-aware Network for 3D Visual Grounding</i>	
Zehan Wang, Haifeng Huang, Yang Zhao, Linjun Li, Xize Cheng, Yichen Zhu, Aoxiong Yin and Zhou Zhao	10612
<i>Goal-Driven Explainable Clustering via Language Descriptions</i>	
Zihan Wang, Jingbo Shang and Ruiqi Zhong	10626

<i>Cross-Lingual Consistency of Factual Knowledge in Multilingual Language Models</i>	
Jirui Qi, Raquel Fernández and Arianna Bisazza	10650
<i>Learning from Mistakes via Cooperative Study Assistant for Large Language Models</i>	
Danqing Wang and Lei Li	10667
<i>Bridging the Digital Divide: Performance Variation across Socio-Economic Factors in Vision-Language Models</i>	
Joan Nwatu, Oana Ignat and Rada Mihalcea	10686
<i>Conceptor-Aided Debiasing of Large Language Models</i>	
Li S. Yifei, Lyle Ungar and João Sedoc	10703
<i>AMR Parsing is Far from Solved: GRAPES, the Granular AMR Parsing Evaluation Suite</i>	
Jonas Groschwitz, Shay B Cohen, Lucia Donatelli and Meaghan Fowlie	10728
<i>Rethinking and Improving Multi-task Learning for End-to-end Speech Translation</i>	
Yuhao Zhang, Chen Xu, Bei Li, Hao Chen, Tong Xiao, Chunliang Zhang and Jingbo Zhu	10753
<i>AD-NLP: A Benchmark for Anomaly Detection in Natural Language Processing</i>	
Matei Bejan, Andrei Manolache and Marius Popescu	10766
<i>Enhancing the Ranking Context of Dense Retrieval through Reciprocal Nearest Neighbors</i>	
George Zerveas, Navid Rekabsaz and Carsten Eickhoff	10779
<i>Cross-Lingual Cross-Target Stance Detection with Dual Knowledge Distillation Framework</i>	
Ruike Zhang, Hanxuan Yang and Wenji Mao	10804
<i>PRESTO: A Multilingual Dataset for Parsing Realistic Task-Oriented Dialogs</i>	
Rahul Goel, Waleed Ammar, Aditya Gupta, Siddharth Vashishtha, Motoki Sano, Faiz Surani, Max Chang, HyunJeong Choe, David Greene, Chuan He, Rattima Nitisaroj, Anna Trukhina, Shachi Paul, Pararth Shah, Rushin Shah and Zhou Yu	10820
<i>An Iteratively Parallel Generation Method with the Pre-Filling Strategy for Document-level Event Extraction</i>	
Guanhua Huang, Runxin Xu, Ying Zeng, Jiaze Chen, Zhouwang Yang and Weinan E	10834
<i>CoMPosT: Characterizing and Evaluating Caricature in LLM Simulations</i>	
Myra Cheng, Tiziano Piccardi and Diyi Yang	10853
<i>Reduce Human Labor On Evaluating Conversational Information Retrieval System: A Human-Machine Collaboration Approach</i>	
Chen Huang, Peixin Qin, Wenqiang Lei and Jiancheng Lv	10876
<i>BERTie Bott's Every Flavor Labels: A Tasty Introduction to Semantic Role Labeling for Galician</i>	
Micaella Bruton and Meriem Beloucif	10892
<i>Program Translation via Code Distillation</i>	
Yufan Huang, Mengnan Qi, Yongqiang Yao, Maoquan Wang, Bin Gu, Colin Clement and Neel Sundaresan	10903
<i>FaMeSumm: Investigating and Improving Faithfulness of Medical Summarization</i>	
Nan Zhang, Yusen Zhang, Wu Guo, Prasenjit Mitra and Rui Zhang	10915
<i>Grammar-Constrained Decoding for Structured NLP Tasks without Finetuning</i>	
Saibo Geng, Martin Josifoski, Maxime Peyrard and Robert West	10932

<i>Systematic word meta-sense extension</i>	
Lei Yu	10953
<i>Evaluating Evaluation Metrics: A Framework for Analyzing NLG Evaluation Metrics using Measurement Theory</i>	
Ziang Xiao, Susu Zhang, Vivian Lai and Q. Vera Liao	10967
<i>Revisiting the Knowledge Injection Frameworks</i>	
Peng Fu, Yiming Zhang, Haobo Wang, Weikang Qiu and Junbo Zhao	10983
<i>We Are What We Repeatedly Do: Inducing and Deploying Habitual Schemas in Persona-Based Responses</i>	
Benjamin Kane and Lenhart K. Schubert	10998
<i>Zero-shot Faithfulness Evaluation for Text Summarization with Foundation Language Model</i>	
Qi Jia, Siyu Ren, Yizhu Liu and Kenny Q. Zhu	11017
<i>TaskWeb: Selecting Better Source Tasks for Multi-task NLP</i>	
Joongwon Kim, Akari Asai, Gabriel Ilharco and Hannaneh Hajishirzi	11032
<i>Improving Bias Mitigation through Bias Experts in Natural Language Understanding</i>	
Eojin Jeon, Mingyu Lee, Juhyeong Park, Yeachan Kim, Wing-Lam Mok and SangKeun Lee	11053
<i>Semi-supervised multimodal coreference resolution in image narrations</i>	
Arushi Goel, Basura Fernando, Frank Keller and Hakan Bilen	11067
<i>A Predictive Factor Analysis of Social Biases and Task-Performance in Pretrained Masked Language Models</i>	
Yi Zhou, Jose Camacho-Collados and Danushka Bollegala	11082
<i>Argument-based Detection and Classification of Fallacies in Political Debates</i>	
Pierpaolo Goffredo, Mariana Chaves Espinoza, Serena Villata and Elena Cabrio	11101
<i>Collaborative Generative AI: Integrating GPT-k for Efficient Editing in Text-to-Image Generation</i>	
Wanrong Zhu, Xinyi Wang, Yujie Lu, Tsu-Jui Fu, Xin Eric Wang, Miguel Eckstein and William Yang Wang	11113
<i>SpEL: Structured Prediction for Entity Linking</i>	
Hassan Shavarani and Anoop Sarkar	11123
<i>Architectural Sweet Spots for Modeling Human Label Variation by the Example of Argument Quality: It's Best to Relate Perspectives!</i>	
Philipp Heinisch, Matthias Orlikowski, Julia Romberg and Philipp Cimiano	11138
<i>Explicit Planning Helps Language Models in Logical Reasoning</i>	
Hongyu Zhao, Kangrui Wang, Mo Yu and Hongyuan Mei	11155
<i>clembench: Using Game Play to Evaluate Chat-Optimized Language Models as Conversational Agents</i>	
Kranti Chalamalasetti, Jana Götze, Sherzod Hakimov, Brielen Madureira, Philipp Sadler and David Schlangen	11174
<i>Explaining with Contrastive Phrasal Highlighting: A Case Study in Assisting Humans to Detect Translation Differences</i>	
Eleftheria Briakou, Navita Goyal and Marine Carpuat	11220
<i>Polyglot or Not? Measuring Multilingual Encyclopedic Knowledge in Foundation Models</i>	
Tim Schott, Daniel Ryan Furman and Shreshta Bhat	11238

<i>Anchoring Fine-tuning of Sentence Transformer with Semantic Label Information for Efficient Truly Few-shot Classification</i>	
Amalie Brogaard Pauli, Leon Derczynski and Ira Assent	11254
<i>UDAPDR: Unsupervised Domain Adaptation via LLM Prompting and Distillation of Rerankers</i>	
Jon Saad-Falcon, Omar Khattab, Keshav Santhanam, Radu Florian, Martin Franz, Salim Roukos, Avirup Sil, Md Arafat Sultan and Christopher Potts	11265
<i>TATA: Stance Detection via Topic-Agnostic and Topic-Aware Embeddings</i>	
Hans William Alexander Hanley and Zakir Durumeric	11280
<i>Data Similarity is Not Enough to Explain Language Model Performance</i>	
Gregory Yauney, Emily Reif and David Mimno	11295
<i>Zero-shot Sharpness-Aware Quantization for Pre-trained Language Models</i>	
Miaoxi Zhu, Qihuang Zhong, Li Shen, Liang Ding, Juhua Liu, Bo Du and Dacheng Tao . .	11305
<i>Deciphering Stereotypes in Pre-Trained Language Models</i>	
Weicheng Ma, Henry Scheible, Brian C Wang, Goutham Veeramachaneni, Pratim Chowdhary, Alan Sun, Andrew Koulogeorge, Lili Wang, Diyi Yang and Soroush Vosoughi	11328
<i>An Integrative Survey on Mental Health Conversational Agents to Bridge Computer Science and Medical Perspectives"</i>	
Young Min Cho, Sunny Rai, Lyle Ungar, João Sedoc and Sharath Chandra Guntuku	11346
<i>Do LLMs Understand Social Knowledge? Evaluating the Sociability of Large Language Models with SocKET Benchmark</i>	
Minje Choi, Jiaxin Pei, Sagar Kumar, Chang Shu and David Jurgens	11370
<i>Interventional Rationalization</i>	
Linan Yue, Qi Liu, Li Wang, Yanqing An, Yichao Du and Zhenya Huang	11404
<i>Don't Take This Out of Context!: On the Need for Contextual Models and Evaluations for Stylistic Rewriting</i>	
Akhila Yerukola, Xuhui Zhou, Elizabeth Clark and Maarten Sap	11419
<i>Axiomatic Preference Modeling for Longform Question Answering</i>	
Corby Rosset, Guoqing Zheng, Victor Dibia, Ahmed Hassan Awadallah and Paul N. Bennett	11445
<i>Countering Misinformation via Emotional Response Generation</i>	
Daniel Russo, Shane Peter Kaszefski-Yaschuk, Jacopo Staiano and Marco Guerini	11476
<i>Seq2seq is All You Need for Coreference Resolution</i>	
Wenzheng Zhang, Sam Wiseman and Karl Stratos	11493
<i>Integrating Language Models into Direct Speech Translation: An Inference-Time Solution to Control Gender Inflection</i>	
Dennis Fucci, Marco Gaido, Sara Papi, Mauro Cettolo, Matteo Negri and Luisa Bentivogli	11505
<i>StoryAnalogy: Deriving Story-level Analogies from Large Language Models to Unlock Analogical Understanding</i>	
Cheng Jiayang, Lin Qiu, Tsz Ho Chan, Tianqing Fang, Weiqi Wang, Chunkit Chan, Dongyu Ru, Qipeng Guo, Hongming Zhang, Yangqiu Song, Yue Zhang and Zheng Zhang	11518
<i>Beyond Detection: A Defend-and-Summarize Strategy for Robust and Interpretable Rumor Analysis on Social Media</i>	
Yi-Ting Chang, Yun-Zhu Song, Yi-Syuan Chen and Hong-Han Shuai	11538

<i>Crystal: Introspective Reasoners Reinforced with Self-Feedback</i>	
Jiacheng Liu, Ramakanth Pasunuru, Hannaneh Hajishirzi, Yejin Choi and Asli Celikyilmaz	11557
<i>DiffS2UT: A Semantic Preserving Diffusion Model for Textless Direct Speech-to-Speech Translation</i>	
Yongxin Zhu, Zhujin Gao, Xinyuan Zhou, Ye Zhongyi and Linli Xu	11573
<i>BioFEG: Generate Latent Features for Biomedical Entity Linking</i>	
Xuhui Sui, Ying Zhang, Xiangrui Cai, Kehui Song, Baohang Zhou, Xiaojie Yuan and Wensheng Zhang	11584
<i>TRIGO: Benchmarking Formal Mathematical Proof Reduction for Generative Language Models</i>	
Jing Xiong, Jianhao Shen, Ye Yuan, Haiming Wang, Yichun Yin, Zhengying Liu, Lin Li, Zhijiang Guo, Qingxing Cao, Yinya Huang, Chuanyang Zheng, Xiaodan Liang, Ming Zhang and Qun Liu	11594
<i>Physician Detection of Clinical Harm in Machine Translation: Quality Estimation Aids in Reliance and Backtranslation Identifies Critical Errors</i>	
Nikita Mehandru, Sweta Agrawal, Yimin Xiao, Ge Gao, Elaine C Khoong, Marine Carpuat and Niloufar Salehi	11633
<i>Vicarious Offense and Noise Audit of Offensive Speech Classifiers: Unifying Human and Machine Disagreement on What is Offensive</i>	
Tharindu Cyril Weerasooriya, Sujana Dutta, Tharindu Ranasinghe, Marcos Zampieri, Christopher M Homan and Ashiqur R. KhudaBukhs	11648
<i>Generating Summaries with Controllable Readability Levels</i>	
Leonardo F. R. Ribeiro, Mohit Bansal and Markus Dreyer	11669
<i>mAggretriever: A Simple yet Effective Approach to Zero-Shot Multilingual Dense Retrieval</i>	
Sheng-Chieh Lin, Amin Ahmad and Jimmy Lin	11688
<i>CodeFusion: A Pre-trained Diffusion Model for Code Generation</i>	
Mukul Singh, José Cambronero, Sumit Gulwani, Vu Le, Carina Suzana Negreanu and Gust Verbruggen	11697
<i>CESAR: Automatic Induction of Compositional Instructions for Multi-turn Dialogs</i>	
Taha Aksu, Devamanyu Hazarika, Shikib Mehri, Seokhwan Kim, Dilek Hakkani-Tur, Yang Liu and Mahdi Namazifar	11709
<i>VECHR: A Dataset for Explainable and Robust Classification of Vulnerability Type in the European Court of Human Rights</i>	
Shanshan Xu, Leon Staufer, Santosh T.Y.S.S, Oana Ichim, Corina Heri and Matthias Grabmair	11738
<i>ACQUIRED: A Dataset for Answering Counterfactual Questions In Real-Life Videos</i>	
Te-Lin Wu, Zi-Yi Dou, Qingyuan Hu, Yu Hou, Nischal Reddy Chandra, Marjorie Freedman, Ralph M. Weischedel and Nanyun Peng	11753
<i>From Parse-Execute to Parse-Execute-Refine: Improving Semantic Parser for Complex Question Answering over Knowledge Base</i>	
Wangzhen Guo, Linyin Luo, Hanjiang Lai and Jian Yin	11771
<i>Reward-Augmented Decoding: Efficient Controlled Text Generation With a Unidirectional Reward Model</i>	
Haikang Deng and Colin Raffel	11781

<i>CORE: A Few-Shot Company Relation Classification Dataset for Robust Domain Adaptation.</i>	
Philipp Borchert, Jochen De Weerd, Kristof Coussement, Arno De Caigny and Marie-Francine Moens	11792
<i>Models See Hallucinations: Evaluating the Factuality in Video Captioning</i>	
Hui Liu and Xiaojun Wan	11807
<i>Back Transcription as a Method for Evaluating Robustness of Natural Language Understanding Models to Speech Recognition Errors</i>	
Marek Kubis, Paweł Marek Skórzewski, Marcin Sowański and Tomasz Ziętkiewicz	11824
<i>Cabbage Sweeter than Cake? Analysing the Potential of Large Language Models for Learning Conceptual Spaces</i>	
Usashi Chatterjee, Amit Gajbhiye and Steven Schockaert	11836
<i>Can Language Models Understand Physical Concepts?</i>	
Lei Li, Jingjing Xu, Qingxiu Dong, Ce Zheng, Xu Sun, Lingpeng Kong and Qi Liu	11843
<i>SPT: Learning to Selectively Insert Prompts for Better Prompt Tuning</i>	
Wei Zhu and Ming Tan	11862
<i>Once Upon a Time in Graph: Relative-Time Pretraining for Complex Temporal Reasoning</i>	
Sen Yang, Xin Li, Lidong Bing and Wai Lam	11879
<i>Expository Text Generation: Imitate, Retrieve, Paraphrase</i>	
Nishant Balepur, Jie Huang and Kevin Chang	11896
<i>Large-scale similarity search with Optimal Transport</i>	
Cléa Mehnia Laouar, Yuki Takezawa and Makoto Yamada	11920
<i>Enhancing Textbooks with Visuals from the Web for Improved Learning</i>	
Janvijay Singh, Vilém Zouhar and Mrinmaya Sachan	11931
<i>Continual Event Extraction with Semantic Confusion Rectification</i>	
Zitao Wang, Xinyi Wang and Wei Hu	11945
<i>An Empirical Study of Translation Hypothesis Ensembling with Large Language Models</i>	
António Farinhas, José G. C. de Souza and Andre Martins	11956
<i>FedTherapist: Mental Health Monitoring with User-Generated Linguistic Expressions on Smartphones via Federated Learning</i>	
Jaemin Shin, Hyungjun Yoon, Seungjoo Lee, Sungjoon Park, Yunxin Liu, Jinho D. Choi and Sung-Ju Lee	11971
<i>Visually-Situated Natural Language Understanding with Contrastive Reading Model and Frozen Large Language Models</i>	
Geewook Kim, Hodong Lee, Daehee Kim, Haeji Jung, Sanghee Park, Yoonsik Kim, Sangdoo Yun, Taeho Kil, Bado Lee and Seunghyun Park	11989
<i>Continual Learning for Multilingual Neural Machine Translation via Dual Importance-based Model Division</i>	
Junpeng Liu, Kaiyu Huang, Hao Yu, Jiuyi Li, Jinsong Su and Degen Huang	12011
<i>SimCSE++: Improving Contrastive Learning for Sentence Embeddings from Two Perspectives</i>	
Jiahao Xu, Wei Shao, Lihui Chen and Lemao Liu	12028
<i>Unlearn What You Want to Forget: Efficient Unlearning for LLMs</i>	
Jiaao Chen and Diyi Yang	12041

<i>Simplicity Level Estimate (SLE): A Learned Reference-Less Metric for Sentence Simplification</i>	
Liam Cripwell, Joël Legrand and Claire Gardent	12053
<i>Precedent-Enhanced Legal Judgment Prediction with LLM and Domain-Model Collaboration</i>	
Yiquan Wu, Siying Zhou, Yifei Liu, Weiming Lu, Xiaozhong Liu, Yating Zhang, Changlong Sun, Fei Wu and Kun Kuang	12060
<i>FACTScore: Fine-grained Atomic Evaluation of Factual Precision in Long Form Text Generation</i>	
Sewon Min, Kalpesh Krishna, Xinxu Lyu, Mike Lewis, Wen-tau Yih, Pang Wei Koh, Mohit Iyyer, Luke Zettlemoyer and Hannaneh Hajishirzi	12076
<i>Calc-X and Calcformers: Empowering Arithmetical Chain-of-Thought through Interaction with Symbolic Systems</i>	
Marek Kadlčík, Michal Štefánik, Ondrej Sotolar and Vlastimil Martinek	12101
<i>CoF-CoT: Enhancing Large Language Models with Coarse-to-Fine Chain-of-Thought Prompting for Multi-domain NLU Tasks</i>	
Hoang H Nguyen, Ye Liu, Chenwei Zhang, Tao Zhang and Philip S. Yu	12109
<i>When Language Models Fall in Love: Animacy Processing in Transformer Language Models</i>	
Michael Hanna, Yonatan Belinkov and Sandro Pezzelle	12120
<i>Improving Unsupervised Relation Extraction by Augmenting Diverse Sentence Pairs</i>	
Qing Wang, Kang Zhou, Qiao Qiao, Yuepei Li and Qi Li	12136
<i>Paraphrase Types for Generation and Detection</i>	
Jan Philip Wahle, Bela Gipp and Terry Ruas	12148
<i>Target-to-Source Augmentation for Aspect Sentiment Triplet Extraction</i>	
Yice Zhang, Yifan Yang, Meng Li, Bin Liang, Shiwei Chen and Ruifeng Xu	12165
<i>PAC-tuning: Fine-tuning Pre-trained Language Models with PAC-driven Perturbed Gradient Descent</i>	
Guangliang Liu, Zhiyu Xue, Xitong Zhang, Kristen Johnson and Rongrong Wang	12178
<i>Emergence of Abstract State Representations in Embodied Sequence Modeling</i>	
Tian Yun, Zilai Zeng, Kunal Handa, Ashish V Thapliyal, Bo Pang, Ellie Pavlick and Chen Sun	12190
<i>Accelerating Toeplitz Neural Network with Constant-time Inference Complexity</i>	
Zhen Qin and Yiran Zhong	12206
<i>Dissecting Recall of Factual Associations in Auto-Regressive Language Models</i>	
Mor Geva, Jasmijn Bastings, Katja Filippova and Amir Globerson	12216
<i>StereoMap: Quantifying the Awareness of Human-like Stereotypes in Large Language Models</i>	
Sullam Jeoung, Yubin Ge and Jana Diesner	12236
<i>Select, Prompt, Filter: Distilling Large Language Models for Summarizing Conversations</i>	
Minh-Quang Pham, Sathish Reddy Indurthi, Shamil Chollampatt and Marco Turchi	12257
<i>Human Raters Cannot Distinguish English Translations from Original English Texts</i>	
Shira Wein	12266
<i>Impressions: Visual Semiotics and Aesthetic Impact Understanding</i>	
Julia Kruk, Caleb Ziems and Diyi Yang	12273

<i>DNA: Denoised Neighborhood Aggregation for Fine-grained Category Discovery</i>	
Wenbin An, Feng Tian, Wenkai Shi, Yan Chen, Qinghua Zheng, QianYing Wang and Ping Chen	
12292	
<i>Prompt as Triggers for Backdoor Attack: Examining the Vulnerability in Language Models</i>	
Shuai Zhao, Jinming Wen, Anh Tuan Luu, Junbo Zhao and Jie Fu	12303
<i>UPRISE: Universal Prompt Retrieval for Improving Zero-Shot Evaluation</i>	
Daixuan Cheng, Shaohan Huang, Junyu Bi, Yuefeng Zhan, Jianfeng Liu, Yujing Wang, Hao Sun, Furu Wei, Weiwei Deng and Qi Zhang	12318
<i>KRLS: Improving End-to-End Response Generation in Task Oriented Dialog with Reinforced Keywords Learning</i>	
Xiao Yu, Qingyang Wu, Kun Qian and Zhou Yu	12338
<i>Large Language Models Only Pass Primary School Exams in Indonesia: A Comprehensive Test on IndoMMLU</i>	
Fajri Koto, Nurul Aisyah, Haonan Li and Timothy Baldwin	12359
<i>Let's Sample Step by Step: Adaptive-Consistency for Efficient Reasoning and Coding with LLMs</i>	
Pranjal Aggarwal, Aman Madaan, Yiming Yang and Mausam	12375
<i>Bridging Information-Theoretic and Geometric Compression in Language Models</i>	
Emily Cheng, Corentin Kervadec and Marco Baroni	12397
<i>Pre-training Language Models for Comparative Reasoning</i>	
Mengxia Yu, Zhihan Zhang, Wenhao Yu and Meng Jiang	12421
<i>Improved Pseudo Data for Machine Translation Quality Estimation with Constrained Beam Search</i>	
Xiang Geng, Yu Zhang, Zhejian Lai, Shuaijie She, Wei Zou, Shimin Tao, Hao Yang, Jiajun Chen and Shujian Huang	12434
<i>Text Embeddings Reveal (Almost) As Much As Text</i>	
John Xavier Morris, Volodymyr Kuleshov, Vitaly Shmatikov and Alexander M Rush	12448
<i>AutoTrial: Prompting Language Models for Clinical Trial Design</i>	
Zifeng Wang, Cao Xiao and Jimeng Sun	12461
<i>Faster Minimum Bayes Risk Decoding with Confidence-based Pruning</i>	
Julius Cheng and Andreas Vlachos	12473
<i>Enhancing Generative Retrieval with Reinforcement Learning from Relevance Feedback</i>	
Yujia Zhou, Zhicheng Dou and Ji-Rong Wen	12481
<i>Multi-Source Probing for Open-Domain Conversational Understanding</i>	
Yuanxi Li, Hao Zhou, Jie Zhou and Minlie Huang	12491
<i>Hallucination Mitigation in Natural Language Generation from Large-Scale Open-Domain Knowledge Graphs</i>	
Xiao Shi, Zhengyuan Zhu, Zeyu Zhang and Chengkai Li	12506
<i>Multi-Source Multi-Type Knowledge Exploration and Exploitation for Dialogue Generation</i>	
Xuanfan Ni, Hongliang Dai, Zhaochun Ren and Piji Li	12522
<i>Focus Your Attention (with Adaptive IIR Filters)</i>	
Shahar Lutati, Itamar Zimerman and Lior Wolf	12538

<i>Identifying Statements Crucial for Awareness of Interpretive Nonsense to Prevent Communication Breakdowns</i>	
Tomoyuki Maekawa and Michita Imai	12550
<i>Multilingual Large Language Models Are Not (Yet) Code-Switchers</i>	
Ruochen Zhang, Samuel Cahyawijaya, Jan Christian Blaise Cruz, Genta Indra Winata and Alham Fikri Aji	12567
<i>Reinforced Target-driven Conversational Promotion</i>	
Huy Quang Dao, Lizi Liao, Dung D. Le and Yuxiang Nie	12583
<i>Identification of Multimodal Stance Towards Frames of Communication</i>	
Maxwell Weinzierl and Sanda Harabagiu	12597
<i>Unsupervised Sounding Pixel Learning</i>	
Yining Zhang, Yanli Ji and Yang Yang	12610
<i>LM vs LM: Detecting Factual Errors via Cross Examination</i>	
Roi Cohen, May Hamri, Mor Geva and Amir Globerson	12621
<i>Large Language Models: The Need for Nuance in Current Debates and a Pragmatic Perspective on Understanding</i>	
Bram van Dijk, Tom Kouwenhoven, Marco Spruit and Max Johannes van Duijn	12641
<i>PIEClass: Weakly-Supervised Text Classification with Prompting and Noise-Robust Iterative Ensemble Training</i>	
Yunyi Zhang, Minhao Jiang, Yu Meng, Yu Zhang and Jiawei Han	12655
<i>MeaeQ: Mount Model Extraction Attacks with Efficient Queries</i>	
Chengwei Dai, Minxuan Lv, Kun Li and Wei Zhou	12671
<i>The CoT Collection: Improving Zero-shot and Few-shot Learning of Language Models via Chain-of-Thought Fine-Tuning</i>	
Seungone Kim, Se June Joo, Doyoung Kim, Joel Jang, Seonghyeon Ye, Jamin Shin and Minjoon Seo	12685
<i>Explaining Interactions Between Text Spans</i>	
Sagnik Ray Choudhury, Pepa Atanasova and Isabelle Augenstein	12709
<i>Predictive Chemistry Augmented with Text Retrieval</i>	
Yujie Qian, Zhening Li, Zhengkai Tu, Connor W. Coley and Regina Barzilay	12731
<i>System Combination via Quality Estimation for Grammatical Error Correction</i>	
Muhammad Reza Qorib and Hwee Tou Ng	12746
<i>Rethinking Negative Pairs in Code Search</i>	
Haochen Li, Xin Zhou, Anh Tuan Luu and Chunyan Miao	12760
<i>Question Answering as Programming for Solving Time-Sensitive Questions</i>	
Xinyu Zhu, Cheng Yang, Bei Chen, Siheng Li, Jian-Guang Lou and Yujiu Yang	12775
<i>Joint Geometrical and Statistical Domain Adaptation for Cross-domain Code Vulnerability Detection</i>	
Qianjin Du, Shiji Zhou, Xiaohui Kuang, Gang Zhao and Jidong Zhai	12791
<i>Revisiting Sparse Retrieval for Few-shot Entity Linking</i>	
Yulin Chen, Zhenran Xu, Baotian Hu and Min Zhang	12801

<i>Controlling Pre-trained Language Models for Grade-Specific Text Simplification</i>	
Sweta Agrawal and Marine Carpuat	12807
<i>CLEVR-Implicit: A Diagnostic Dataset for Implicit Reasoning in Referring Expression Comprehension</i>	
Jingwei Zhang, Xin Wu and Yi Cai	12820
<i>“Are Your Explanations Reliable?” Investigating the Stability of LIME in Explaining Text Classifiers by Marrying XAI and Adversarial Attack</i>	
Christopher Burger, Lingwei Chen and Thai Le	12831
<i>CQE: A Comprehensive Quantity Extractor</i>	
Satya Almasian, Vivian Kazakova, Philipp Göldner and Michael Gertz	12845
<i>Context Compression for Auto-regressive Transformers with Sentinel Tokens</i>	
Siyu Ren, Qi Jia and Kenny Q. Zhu	12860
<i>A Unified View of Evaluation Metrics for Structured Prediction</i>	
Yunmo Chen, William Gantt, Tongfei Chen, Aaron Steven White and Benjamin Van Durme	12868
<i>A Deeper (Autoregressive) Approach to Non-Convergent Discourse Parsing</i>	
Oren Tsur and Yoav Tulpan	12883
<i>We are Who We Cite: Bridges of Influence Between Natural Language Processing and Other Academic Fields</i>	
Jan Philip Wahle, Terry Ruas, Mohamed Abdalla, Bela Gipp and Saif M. Mohammad	12896
<i>Ties Matter: Meta-Evaluating Modern Metrics with Pairwise Accuracy and Tie Calibration</i>	
Daniel Deutsch, George Foster and Markus Freitag	12914
<i>SODA: Million-scale Dialogue Distillation with Social Commonsense Contextualization</i>	
Hyunwoo Kim, Jack Hessel, Liwei Jiang, Peter West, Ximing Lu, Youngjae Yu, Pei Zhou, Ronan Le Bras, Malihe Alikhani, Gunhee Kim, Maarten Sap and Yejin Choi	12930
<i>Multi-view Contrastive Learning for Entity Typing over Knowledge Graphs</i>	
Zhiwei Hu, Victor Gutierrez Basulto, Zhiliang Xiang, Ru Li and Jeff Z. Pan	12950
<i>MailEx: Email Event and Argument Extraction</i>	
Saurabh Srivastava, Gaurav Singh, Shou Matsumoto, Ali K Raz, Paulo Costa, Joshua Campbell Poore and Ziyu Yao	12964
<i>Optimized Tokenization for Transcribed Error Correction</i>	
Tomer Wullach and Shlomo Chazan	12988
<i>Beware of Model Collapse! Fast and Stable Test-time Adaptation for Robust Question Answering</i>	
Yi Su, Yixin Ji, Juntao Li, Hai Ye and Min Zhang	12998
<i>Generative Adversarial Training with Perturbed Token Detection for Model Robustness</i>	
Jiahao Zhao and Wenji Mao	13012
<i>Multi-Task Knowledge Distillation with Embedding Constraints for Scholarly Keyphrase Boundary Classification</i>	
Seo Yeon Park and Cornelia Caragea	13026
<i>Set Learning for Generative Information Extraction</i>	
Jiangnan Li, Yice Zhang, Bin Liang, Kam-Fai Wong and Ruifeng Xu	13043
<i>Large Language Models and Multimodal Retrieval for Visual Word Sense Disambiguation</i>	
Anastasia Kritharoula, Maria Lymperaiou and Giorgos Stamou	13053

<i>Be Selfish, But Wisely: Investigating the Impact of Agent Personality in Mixed-Motive Human-Agent Interactions</i>	
Kushal Chawla, Ian Wu, Yu Rong, Gale Lucas and Jonathan Gratch	13078
<i>Doolittle: Benchmarks and Corpora for Academic Writing Formalization</i>	
Shizhe Diao, Yongyu Lei, Liangming Pan, Tianqing Fang, Wangchunshu Zhou, Sedrick Scott Keh, Min-Yen Kan and Tong Zhang	13093
<i>Token Prediction as Implicit Classification to Identify LLM-Generated Text</i>	
Yutian Chen, Hao Kang, Vivian Zhai, Liangze Li, Rita Singh and Bhiksha Raj	13112
<i>On Evaluation of Bangla Word Analogies</i>	
Mousumi Akter, Souvika Sarkar and Shubhra Kanti Karmaker Santu	13121
<i>Reconstruct Before Summarize: An Efficient Two-Step Framework for Condensing and Summarizing Meeting Transcripts</i>	
Haochen Tan, Han Wu, Wei Shao, Xinyun Zhang, Mingjie Zhan, Zhaohui Hou, Ding Liang and Linqi Song	13128
<i>XLM-V: Overcoming the Vocabulary Bottleneck in Multilingual Masked Language Models</i>	
Davis Liang, Hila Gonen, Yuning Mao, Rui Hou, Naman Goyal, Marjan Ghazvininejad, Luke Zettlemoyer and Madian Khabsa	13142
<i>Character-LLM: A Trainable Agent for Role-Playing</i>	
Yunfan Shao, Linyang Li, Junqi Dai and Xipeng Qiu	13153
<i>Natural Language Decompositions of Implicit Content Enable Better Text Representations</i>	
Alexander Hoyle, Rupak Sarkar, Pranav Goel and Philip Resnik	13188
<i>A Scalable Framework for Table of Contents Extraction from Complex ESG Annual Reports</i>	
Xinyu Wang, Lin Gui and Yulan He	13215
<i>Semantic Space Grounded Weighted Decoding for Multi-Attribute Controllable Dialogue Generation</i>	
Zhiling Zhang, Mengyue Wu and Kenny Q. Zhu	13230
<i>How do languages influence each other? Studying cross-lingual data sharing during LM fine-tuning</i>	
Rochelle Choenni, Dan Garrette and Ekaterina Shutova	13244
<i>COFFEE: Counterfactual Fairness for Personalized Text Generation in Explainable Recommendation</i>	
Nan Wang, Qifan Wang, Yi-Chia Wang, Maziar Sanjabi, Jingzhou Liu, Hamed Firooz, Hongning Wang and Shaoliang Nie	13258
<i>NameGuess: Column Name Expansion for Tabular Data</i>	
Jiani Zhang, Zhengyuan Shen, Balasubramaniam Srinivasan, Shen Wang, Huzefa Rangwala and George Karypis	13276
<i>BLESS: Benchmarking Large Language Models on Sentence Simplification</i>	
Tannon Kew, Alison Chi, Laura Vásquez-Rodríguez, Sweta Agrawal, Dennis Aumiller, Fernando Alva-Manchego and Matthew Shardlow	13291
<i>To Build Our Future, We Must Know Our Past: Contextualizing Paradigm Shifts in Natural Language Processing</i>	
Sireesh Gururaja, Amanda Bertsch, Clara Na, David Gray Widder and Emma Strubell	13310
<i>PALS: Personalized Active Learning for Subjective Tasks in NLP</i>	
Kamil Kanclerz, Konrad Karanowski, Julita Bielaniec, Marcin Gruz, Piotr Miłkowski, Jan Kocon and Przemysław Kazienko	13326

<i>ViStruct: Visual Structural Knowledge Extraction via Curriculum Guided Code-Vision Representation</i> Yangyi Chen, Xingyao Wang, Manling Li, Derek Hoiem and Heng Ji	13342
<i>LLMLingua: Compressing Prompts for Accelerated Inference of Large Language Models</i> Huiqiang Jiang, Qianhui Wu, Chin-Yew Lin, Yuqing Yang and Lili Qiu.....	13358
<i>EXPLAIN, EDIT, GENERATE: Rationale-Sensitive Counterfactual Data Augmentation for Multi-hop Fact Verification</i> Yingjie Zhu, Jiasheng Si, Yibo Zhao, Haiyang Zhu, Deyu Zhou and Yulan He	13377
<i>An Exploration of Left-Corner Transformations</i> Andreas Opedal, Eleftheria Tsipidi, Tiago Pimentel, Ryan Cotterell and Tim Vieira.....	13393
<i>Characterizing and Verifying Scientific Claims: Qualitative Causal Structure is All You Need</i> Jinxuan Wu, Wenhan Chao, Xian Zhou and Zhunchen Luo.....	13428
<i>FOCUS: Effective Embedding Initialization for Monolingual Specialization of Multilingual Models</i> Konstantin Dobler and Gerard de Melo	13440
<i>ByteSized32: A Corpus and Challenge Task for Generating Task-Specific World Models Expressed as Text Games</i> Ruoyao Wang, Graham Todd, Xingdi Yuan, Ziang Xiao, Marc-Alexandre Côté and Peter Jansen	13455
<i>Skill-Based Few-Shot Selection for In-Context Learning</i> Shengnan An, Bo Zhou, Zeqi Lin, Qiang Fu, Bei Chen, Nanning Zheng, Weizhu Chen and Jian-Guang Lou	13472
<i>MaNtLE: Model-agnostic Natural Language Explainer</i> Rakesh R Menon, Kerem Zaman and Shashank Srivastava	13493
<i>PTP: Boosting Stability and Performance of Prompt Tuning with Perturbation-Based Regularizer</i> Lichang Chen, Jiuhai Chen, Heng Huang and Minhao Cheng	13512
<i>Ling-CL: Understanding NLP Models through Linguistic Curricula</i> Mohamed Elgaar and Hadi Amiri	13526
<i>Towards Unsupervised Recognition of Token-level Semantic Differences in Related Documents</i> Jannis Vamvas and Rico Sennrich	13543
<i>Towards a Better Understanding of Variations in Zero-Shot Neural Machine Translation Performance</i> Shaomu Tan and Christof Monz	13553
<i>SEER : A Knapsack approach to Exemplar Selection for In-Context HybridQA</i> Jonathan Tonglet, Manon Reusens, Philipp Borchert and Bart Baesens	13569
<i>Conversation Chronicles: Towards Diverse Temporal and Relational Dynamics in Multi-Session Conversations</i> Jihyoung Jang, Minseong Boo and Hyoungun Kim	13584
<i>DueT: Image-Text Contrastive Transfer Learning with Dual-adapter Tuning</i> Taku Hasegawa, Kyosuke Nishida, Koki Maeda and Kuniko Saito.....	13607
<i>Towards a Unified Conversational Recommendation System: Multi-task Learning via Contextualized Knowledge Distillation</i> Yeongseo Jung, Eunseo Jung and Lei Chen	13625

<i>CLAIR: Evaluating Image Captions with Large Language Models</i>	
David Chan, Suzanne Petryk, Joseph E. Gonzalez, Trevor Darrell and John Canny	13638
<i>MoPe: Model Perturbation based Privacy Attacks on Language Models</i>	
Marvin Li, Jason Wang, Jeffrey George Wang and Seth Neel	13647
<i>q2d: Turning Questions into Dialogs to Teach Models How to Search</i>	
Yonatan Bitton, Shlomi Cohen-Ganor, Ido Hakimi, Yoad Lewenberg, Roei Aharoni and Enav Weinreb	13661
<i>Aligning Large Language Models through Synthetic Feedback</i>	
Sungdong Kim, Sanghwan Bae, Jamin Shin, Soyoung Kang, Donghyun Kwak, Kang Min Yoo and Minjoon Seo	13677
<i>You Told Me That Joke Twice: A Systematic Investigation of Transferability and Robustness of Humor Detection Models</i>	
Alexander Baranov, Vladimir Kniazhevsky and Pavel Braslavski	13701
<i>Reading Order Matters: Information Extraction from Visually-rich Documents by Token Path Prediction</i>	
Chong Zhang, Ya Guo, Yi Tu, Huan Chen, Jinyang Tang, Huijia Zhu, Qi Zhang and Tao Gui	13716
<i>Empower Nested Boolean Logic via Self-Supervised Curriculum Learning</i>	
Hongqiu Wu, Linfeng Liu, Hai Zhao and Min Zhang	13731
<i>The Sentiment Problem: A Critical Survey towards Deconstructing Sentiment Analysis</i>	
Pranav Narayanan Venkit, Mukund Srinath, Sanjana Gautam, Saranya Venkatraman, Vipul Gupta, Rebecca J. Passonneau and Shomir Wilson	13743
<i>Poisoning Retrieval Corpora by Injecting Adversarial Passages</i>	
Zexuan Zhong, Ziqing Huang, Alexander Wettig and Danqi Chen	13764
<i>DADA: Dialect Adaptation via Dynamic Aggregation of Linguistic Rules</i>	
Yanchen Liu, William Barr Held and Diyi Yang	13776
<i>Clustering Pseudo Language Family in Multilingual Translation Models with Fisher Information Matrix</i>	
Xinyu Ma, Xuebo Liu and Min Zhang	13794
<i>Unifying Discrete and Continuous Representations for Unsupervised Paraphrase Generation</i>	
Mingfeng Xue, Dayiheng Liu, Wenqiang Lei, Jie Fu, Jian Lan, Mei Li, Baosong Yang, Jun Xie, Yidan Zhang, Dezhong Peng and Jiancheng Lv	13805
<i>The Benefits of Label-Description Training for Zero-Shot Text Classification</i>	
Lingyu Gao, Debanjan Ghosh and Kevin Gimpel	13823
<i>Multilingual Pixel Representations for Translation and Effective Cross-lingual Transfer</i>	
Elizabeth Salesky, Neha Verma, Philipp Koehn and Matt Post	13845
<i>Finding Authentic Counterhate Arguments: A Case Study with Public Figures</i>	
Abdullah Albanyan, Ahmed Hassan and Eduardo Blanco	13862
<i>Can We Edit Multimodal Large Language Models?</i>	
Siyuan Cheng, Bozhong Tian, Qingbin Liu, Xi Chen, Yongheng Wang, Huajun Chen and Ningyu Zhang	13877
<i>Exploring Discourse Structure in Document-level Machine Translation</i>	
Xinyu Hu and Xiaojun Wan	13889

<i>ClusterLLM: Large Language Models as a Guide for Text Clustering</i>	
Yuwei Zhang, Zihan Wang and Jingbo Shang	13903
<i>CodeBERTScore: Evaluating Code Generation with Pretrained Models of Code</i>	
Shuyan Zhou, Uri Alon, Sumit Agarwal and Graham Neubig	13921
<i>Learn and Consolidate: Continual Adaptation for Zero-Shot and Multilingual Neural Machine Translation</i>	
Kaiyu Huang, Peng Li, Junpeng Liu, Maosong Sun and Yang Liu	13938
<i>e-THERAPIST: I suggest you to cultivate a mindset of positivity and nurture uplifting thoughts</i>	
Kshitij Mishra, Priyanshu Priya, Manisha Burja and Asif Ekbal	13952
<i>AfriSenti: A Twitter Sentiment Analysis Benchmark for African Languages</i>	
Shamsuddeen Hassan Muhammad, Idris Abdulmumin, Abinew Ali Ayele, Nedjma Ousidhoum, David Ifeoluwa Adelani, Seid Muhie Yimam, Ibrahim Said Ahmad, Meriem Beloucif, Saif M. Mohammad, Sebastian Ruder, Oumaima Hourrane, Alipio Jorge, Pavel Brazdil, Felermino D. M. A. Ali, Davis David, Salomey Osei, Bello Shehu-Bello, Falalu Ibrahim Lawan, Tajuddeen Gwadabe, Samuel Rutunda, Tadesse Destaw Belay, Wendimu Baye Messelle, Hailu Beshada Balcha, Sisay Adugna Chala, Hagos Tesfahun Gebremichael, Bernard Opoku and Stephen Arthur	13968
<i>Quantifying Character Similarity with Vision Transformers</i>	
Xinmei Yang, Abhishek Arora, Shao-Yu Jheng and Melissa Dell	13982
<i>Syllogistic Reasoning for Legal Judgment Analysis</i>	
Wentao Deng, Jiahuan Pei, Keyi Kong, Zhe Chen, Furu Wei, Yujun Li, Zhaochun Ren, Zhumin Chen and Pengjie Ren	13997
<i>Improving Transformer-based Program Repair Model through False Behavior Diagnosis</i>	
Youngkyoung Kim, Misoo Kim and Eunseok Lee	14010
<i>SUT: Active Defects Probing for Transcompiler Models</i>	
Mengnan Qi, Yufan Huang, Maoquan Wang, Yongqiang Yao, Zihan Liu, Bin Gu, Colin Clement and Neel Sundaresan	14024
<i>KCTS: Knowledge-Constrained Tree Search Decoding with Token-Level Hallucination Detection</i>	
Sehyun Choi, Tianqing Fang, Zhaowei Wang and Yangqiu Song	14035
<i>CRUSH4SQL: Collective Retrieval Using Schema Hallucination For Text2SQL</i>	
Mayank Kothiyari, Dhruva Dhingra, Sunita Sarawagi and Soumen Chakrabarti	14054
<i>This Reads Like That: Deep Learning for Interpretable Natural Language Processing</i>	
Claudio Fanconi, Moritz Vandenbirtz, Severin Husmann and Julia E Vogt	14067
<i>Incorporating Structured Representations into Pretrained Vision & Language Models Using Scene Graphs</i>	
Roei Herzig, Alon Mendelson, Leonid Karlinsky, Assaf Arbelle, Rogerio Feris, Trevor Darrell and Amir Globerson	14077
<i>TLM: Token-Level Masking for Transformers</i>	
Yangjun Wu, Kebin Fang, Dongxiang Zhang, Han Wang, Hao Zhang and Gang Chen	14099
<i>Addressing NER Annotation Noises with Uncertainty-Guided Tree-Structured CRFs</i>	
Jian Liu, Weichang Liu, Yufeng Chen, Jinan Xu and Zhe Zhao	14112
<i>Hi Guys or Hi Folks? Benchmarking Gender-Neutral Machine Translation with the GeNTE Corpus</i>	
Andrea Piergentili, Beatrice Savoldi, Dennis Fucci, Matteo Negri and Luisa Bentivogli ...	14124

<i>Multilingual Holistic Bias: Extending Descriptors and Patterns to Unveil Demographic Biases in Languages at Scale</i>	
Marta R. Costa-jussà, Pierre Andrews, Eric Michael Smith, Prangthip Hansanti, Christophe Ropers, Elahe Kalbassi, Cynthia Gao, Daniel Edward Licht and Carleigh Wood	14141
<i>GlobalBench: A Benchmark for Global Progress in Natural Language Processing</i>	
Yueqi Song, Simran Khanuja, Pengfei Liu, Fahim Faisal, Alissa Ostapenko, Genta Indra Winata, Alham Fikri Aji, Samuel Cahyawijaya, Yulia Tsvetkov, Antonios Anastasopoulos and Graham Neubig	14157
<i>DetGPT: Detect What You Need via Reasoning</i>	
Renjie Pi, Jiahui Gao, Shizhe Diao, Rui Pan, Hanze Dong, Jipeng Zhang, Lewei Yao, Jianhua Han, Hang Xu, Lingpeng Kong and Tong Zhang	14172
<i>Language Models with Rationality</i>	
Nora Kassner, Oyvind Tafjord, Ashish Sabharwal, Kyle Richardson, Hinrich Schuetze and Peter Clark	14190
<i>Self-Improvement of Non-autoregressive Model via Sequence-Level Distillation</i>	
Yusheng Liao, Shuyang Jiang, Yiqi Li, Yu Wang and Yanfeng Wang	14202
<i>Mitigating Temporal Misalignment by Discarding Outdated Facts</i>	
Michael J.Q. Zhang and Eunsol Choi	14213
<i>Open-world Semi-supervised Generalized Relation Discovery Aligned in a Real-world Setting</i>	
William P Hogan, Jiacheng Li and Jingbo Shang	14227
<i>IEKG: A Commonsense Knowledge Graph for Idiomatic Expressions</i>	
Ziheng Zeng, Kellen Tan Cheng, Srihari Venkat Nanniyur, Jianing Zhou and Suma Bhat . .	14243
<i>Bias Neutralization in Non-Parallel Texts: A Cyclic Approach with Auxiliary Guidance</i>	
Karthic Madanagopal and James Caverlee	14265
<i>Fighting Fire with Fire: The Dual Role of LLMs in Crafting and Detecting Elusive Disinformation</i>	
Jason S Lucas, Adaku Uchendu, Michiharu Yamashita, Jooyoung Lee, Shaurya Rohatgi and Dongwon Lee	14279
<i>SMoP: Towards Efficient and Effective Prompt Tuning with Sparse Mixture-of-Prompts</i>	
Joon-Young Choi, Junho Kim, Jun-Hyung Park, Wing-Lam Mok and SangKeun Lee	14306
<i>BRAINTEASER: Lateral Thinking Puzzles for Large Language Models</i>	
Yifan Jiang, Filip Ilievski, Kaixin Ma and Zhivar Sourati	14317
<i>When are Lemons Purple? The Concept Association Bias of Vision-Language Models</i>	
Yingtian Tang, Yutaro Yamada, Yoyo Minzhi Zhang and Ilker Yildirim	14333
<i>What Comes Next? Evaluating Uncertainty in Neural Text Generators Against Human Production Variability</i>	
Mario Giulianelli, Joris Baan, Wilker Aziz, Raquel Fernández and Barbara Plank	14349
<i>Text Representation Distillation via Information Bottleneck Principle</i>	
Yanzhao Zhang, Dingkun Long, Zehan Li and Pengjun Xie	14372
<i>Let GPT be a Math Tutor: Teaching Math Word Problem Solvers with Customized Exercise Generation</i>	
Zhenwen Liang, Wenhao Yu, Tanmay Rajpurohit, Peter Clark, Xiangliang Zhang and Ashwin Kalyan	14384

<i>FANToM: A Benchmark for Stress-testing Machine Theory of Mind in Interactions</i>	
Hyunwoo Kim, Melanie Sclar, Xuhui Zhou, Ronan Le Bras, Gunhee Kim, Yejin Choi and Maarten Sap	14397
<i>Exploring the Boundaries of GPT-4 in Radiology</i>	
Qianchu Liu, Stephanie Hyland, Shruthi Bannur, Kenza Bouzid, Daniel C. Castro, Maria Teodora Wetscherek, Robert Tinn, Harshita Sharma, Fernando Pérez-García, Anton Schwaighofer, Pranav Rajpurkar, Sameer Tajdin Khanna, Hoifung Poon, Naoto Usuyama, Anja Thieme, Aditya V. Nori, Matthew P. Lungren, Ozan Oktay and Javier Alvarez-Valle	14414
<i>A Frustratingly Easy Post-Training Quantization Scheme for LLMs</i>	
Yongkweon Jeon, Chungman Lee, Kyungphil Park and Ho-young Kim	14446
<i>A Comprehensive Evaluation of Biomedical Entity Linking Models</i>	
David Kartchner, Jennifer Deng, Shubham Lohiya, Tejasri Kopparthi, Prasanth Bathala, Daniel Domingo-Fernández and Cassie S. Mitchell	14462
<i>Exploring Jiu-Jitsu Argumentation for Writing Peer Review Rebuttals</i>	
Sukannya Purkayastha, Anne Lauscher and Iryna Gurevych	14479
<i>LIMIT: Language Identification, Misidentification, and Translation using Hierarchical Models in 350+ Languages</i>	
Milind Agarwal, Md Mahfuz Ibn Alam and Antonios Anastasopoulos	14496
<i>FreeAL: Towards Human-Free Active Learning in the Era of Large Language Models</i>	
Ruixuan Xiao, Yiwen Dong, Junbo Zhao, Runze Wu, Minmin Lin, Gang Chen and Haobo Wang	14520
<i>API-Assisted Code Generation for Question Answering on Varied Table Structures</i>	
Yihan Cao, Shuyi Chen, Ryan Liu, Zhiruo Wang and Daniel Fried	14536
<i>Data Factors for Better Compositional Generalization</i>	
Xiang Zhou, Yichen Jiang and Mohit Bansal	14549
<i>ChatEdit: Towards Multi-turn Interactive Facial Image Editing via Dialogue</i>	
Xing Cui, Zekun Li, Pei Pei Li, Yibo Hu, Hailin Shi, Chunshui Cao and Zhaofeng He	14567
<i>Bridging Continuous and Discrete Spaces: Interpretable Sentence Representation Learning via Compositional Operations</i>	
James Y. Huang, Wenlin Yao, Kaiqiang Song, Hongming Zhang, Muhao Chen and Dong Yu	14584
<i>Outlier Dimensions Encode Task Specific Knowledge</i>	
William Rudman, Catherine Chen and Carsten Eickhoff	14596
<i>Hi-ArG: Exploring the Integration of Hierarchical Argumentation Graphs in Language Pretraining</i>	
Jingcong Liang, Rong Ye, Meng Han, Qi Zhang, Ruofei Lai, Xinyu Zhang, Zhao Cao, Xuanjing Huang and Zhongyu Wei	14606
<i>Biomedical Named Entity Recognition via Dictionary-based Synonym Generalization</i>	
Zihao Fu, Yixuan Su, Zaiqiao Meng and Nigel Collier	14621
<i>GNAT: A General Narrative Alignment Tool</i>	
Tanzir Pial and Steven Skiena	14636
<i>Self-Ensemble of N-best Generation Hypotheses by Lexically Constrained Decoding</i>	
Ryota Miyano, Tomoyuki Kajiwaru and Yuki Arase	14653

<i>UniChart: A Universal Vision-language Pretrained Model for Chart Comprehension and Reasoning</i>	
Ahmed Masry, Parsa Kavehzadeh, Do Xuan Long, Enamul Hoque and Shafiq Joty	14662
<i>Merging Experts into One: Improving Computational Efficiency of Mixture of Experts</i>	
Shwai He, Run-Ze Fan, Liang Ding, Li Shen, Tianyi Zhou and Dacheng Tao	14685
<i>Distance-Based Propagation for Efficient Knowledge Graph Reasoning</i>	
Harry Shomer, Yao Ma, Juanhui Li, Bo Wu, Charu C. Aggarwal and Jiliang Tang	14692
<i>What to Read in a Contract? Party-Specific Summarization of Legal Obligations, Entitlements, and Prohibitions</i>	
Abhilasha Sancheti, Aparna Garimella, Balaji Vasan Srinivasan and Rachel Rudinger	14708
<i>Enhancing Computation Efficiency in Large Language Models through Weight and Activation Quantization</i>	
Janghwan Lee, Minsoo Kim, Seungcheol Baek, Seok Joong Hwang, Wonyong Sung and Jungwook Choi	14726
<i>CP-BCS: Binary Code Summarization Guided by Control Flow Graph and Pseudo Code</i>	
Tong Ye, Lingfei Wu, Tengfei Ma, Xuhong Zhang, Yangkai Du, Peiyu Liu, Shouling Ji and Wenhai Wang	14740
<i>Assessing Step-by-Step Reasoning against Lexical Negation: A Case Study on Syllogism</i>	
Mengyu Ye, Tatsuki Kuribayashi, Jun Suzuki, Goro Kobayashi and Hiroaki Funayama	14753
<i>Chain-of-Thought Tuning: Masked Language Models can also Think Step By Step in Natural Language Understanding</i>	
Caoyun Fan, Jidong Tian, Yitian Li, Wenqing Chen, Hao He and Yaohui Jin	14774
<i>Large Language Models are Complex Table Parsers</i>	
Bowen Zhao, Changkai Ji, Yuejie Zhang, Wen He, Yingwen Wang, Qing Wang, Rui Feng and Xiaobo Zhang	14786
<i>R2H: Building Multimodal Navigation Helpers that Respond to Help Requests</i>	
Yue Fan, Jing Gu, Kaizhi Zheng and Xin Eric Wang	14803
<i>Speech-enriched Memory for Inference-time Adaptation of ASR Models to Word Dictionaries</i>	
Ashish Mittal, Sunita Sarawagi, Preethi Jyothi, George Saon and Gakuto Kurata	14820
<i>Generative Table Pre-training Empowers Models for Tabular Prediction</i>	
Tianping Zhang, Shaowen Wang, Shuicheng Yan, Li Jian and Qian Liu	14836
<i>Learning to Describe for Predicting Zero-shot Drug-Drug Interactions</i>	
Fangqi Zhu, Yongqi Zhang, Lei Chen, Bing Qin and Ruifeng Xu	14855
<i>A Simple Baseline for Knowledge-Based Visual Question Answering</i>	
Alexandros Xenos, Themis Stafylakis, Ioannis Patras and Georgios Tzimiropoulos	14871
<i>Unveiling the Essence of Poetry: Introducing a Comprehensive Dataset and Benchmark for Poem Summarization</i>	
Ridwan Mahbub, Ifrad Towhid Khan, Samiha Shafiq Anuva, Md Shihab Shahriar, Md Tahmid Rahman Laskar and Sabbir Ahmed	14878
<i>Privacy Implications of Retrieval-Based Language Models</i>	
Yangsibo Huang, Samyak Gupta, Zexuan Zhong, Kai Li and Danqi Chen	14887

<i>IMTLab: An Open-Source Platform for Building, Evaluating, and Diagnosing Interactive Machine Translation Systems</i>	
Xu Huang, Zhirui Zhang, Ruize Gao, Yichao Du, Lemao Liu, Guoping Huang, Shuming Shi, Jiajun Chen and Shujian Huang	14903
<i>Is ChatGPT Good at Search? Investigating Large Language Models as Re-Ranking Agents</i>	
Weiwei Sun, Lingyong Yan, Xinyu Ma, Shuaiqiang Wang, Pengjie Ren, Zhumin Chen, Dawei Yin and Zhaochun Ren	14918
<i>DiNeR: A Large Realistic Dataset for Evaluating Compositional Generalization</i>	
Chengang Hu, Xiao Liu and Yansong Feng	14938
<i>Can Pre-trained Vision and Language Models Answer Visual Information-Seeking Questions?</i>	
Yang Chen, Hexiang Hu, Yi Luan, Haitian Sun, Soravit Changpinyo, Alan Ritter and Ming-Wei Chang	14948
<i>EDeR: Towards Understanding Dependency Relations Between Events</i>	
Ruiqi Li, Patrik Haslum and Leyang Cui	14969
<i>It Ain't Over: A Multi-aspect Diverse Math Word Problem Dataset</i>	
Jiwoo Kim, Youngbin Kim, Ilwoong Baek, JinYeong Bak and Jongwuk Lee	14984
<i>Dr ChatGPT tell me what I want to hear: How different prompts impact health answer correctness</i>	
Bevan Koopman and Guido Zuccon	15012
<i>kNN-LM Does Not Improve Open-ended Text Generation</i>	
Shufan Wang, Yixiao Song, Andrew Drozdov, Aparna Garimella, Varun Manjunatha and Mohit Iyyer	15023
<i>Towards A Unified View of Sparse Feed-Forward Network in Pretraining Large Language Model</i>	
Zeyu Liu, Tim Dettmers, Xi Victoria Lin, Veselin Stoyanov and Xian Li	15038
<i>Exploring the Impact of Model Scaling on Parameter-Efficient Tuning</i>	
Yusheng Su, Chi-Min Chan, Jiali Cheng, Yujia Qin, Yankai Lin, Shengding Hu, Zonghan Yang, Ning Ding, Xingzhi Sun, Guotong Xie, Zhiyuan Liu and Maosong Sun	15062
<i>STAIR: Learning Sparse Text and Image Representation in Grounded Tokens</i>	
Chen Chen, Bowen Zhang, Liangliang Cao, Jiguang Shen, Tom Gunter, Albin Madappally Jose, Alexander T Toshev, Yantao Zheng, Jonathon Shlens, Ruoming Pang and Yinfei Yang	15079
<i>Crossing the Threshold: Idiomatic Machine Translation through Retrieval Augmentation and Loss Weighting</i>	
Emmy Liu, Aditi Chaudhary and Graham Neubig	15095
<i>CoRec: An Easy Approach for Coordination Recognition</i>	
Qing Wang, Haojie Jia, Wenfei Song and Qi Li	15112
<i>A linear time approximation of Wasserstein distance with word embedding selection</i>	
Sho Otao and Makoto Yamada	15121
<i>Exchange-of-Thought: Enhancing Large Language Model Capabilities through Cross-Model Communication</i>	
Zhangyue Yin, Qiushi Sun, Cheng Chang, Qipeng Guo, Junqi Dai, Xuanjing Huang and Xipeng Qiu	15135
<i>Conversation Understanding using Relational Temporal Graph Neural Networks with Auxiliary Cross-Modality Interaction</i>	
Cam Van Thi Nguyen, Tuan Anh Mai, Son Le The, Dang Hai Kieu and Duc-Trong Le	15154

<i>Connecting degree and polarity: An artificial language learning study</i>	
Lisa Bylinina, Alexey Tikhonov and Ekaterina Garmash	15168
<i>Prompting with Pseudo-Code Instructions</i>	
Mayank Mishra, Prince Kumar, Riyaz Ahmad Bhat, Rudra Murthy, Danish Contractor and Srikanth G. Tamilselvam	15178
<i>CRAB: Assessing the Strength of Causal Relationships Between Real-world Events</i>	
Angelika Romanou, Syrielle Montariol, Debjit Paul, Leo Laugier, Karl Aberer and Antoine Bosselet	15198
<i>NORMSAGE: Multi-Lingual Multi-Cultural Norm Discovery from Conversations On-the-Fly</i>	
Yi Fung, Tuhin Chakrabarty, Hao Guo, Owen Rambow, Smaranda Muresan and Heng Ji ..	15217
<i>A State-Vector Framework for Dataset Effects</i>	
Esmat Sahak, Zining Zhu and Frank Rudzicz	15231
<i>Challenges in Context-Aware Neural Machine Translation</i>	
Linghao Jin, Jacqueline He, Jonathan May and Xuezhe Ma	15246
<i>Task-Adaptive Tokenization: Enhancing Long-Form Text Generation Efficacy in Mental Health and Beyond</i>	
Siyang Liu, Naihao Deng, Sahand Sabour, Yilin Jia, Minlie Huang and Rada Mihalcea ...	15264
<i>FACTIFY3M: A benchmark for multimodal fact verification with explainability through 5W Question-Answering</i>	
Megha Chakraborty, Khushbu Pahwa, Anku Rani, Shreyas Chatterjee, Dwip Dalal, Harshit Dave, Ritvik G, Preethi Gurumurthy, Adarsh Ashok Mahor, Samahriti Mukherjee, Aditya Pakala, Ishan Paul, Janvita Reddy, Arghya Sarkar, Kinjal Sensharma, Aman Chadha, Amit P. Sheth and Amitava Das	15282
<i>Building Multi-domain Dialog State Trackers from Single-domain Dialogs</i>	
Qi Zhu, Zheng Zhang, Xiaoyan Zhu and Minlie Huang	15323
<i>Specialist or Generalist? Instruction Tuning for Specific NLP Tasks</i>	
Chufan Shi, Yixuan Su, Cheng Yang, Yujiu Yang and Deng Cai	15336
<i>Making Large Language Models Better Data Creators</i>	
Dong-Ho Lee, Jay Pujara, Mohit Sewak, Ryen W White and Sujay Kumar Jauhar	15349
<i>Hallucination Detection for Generative Large Language Models by Bayesian Sequential Estimation</i>	
Xiaohua Wang, Yuliang Yan, Longtao Huang, Xiaoqing Zheng and Xuanjing Huang	15361
<i>Guideline Learning for In-Context Information Extraction</i>	
Chaoxu Pang, Yixuan Cao, Qiang Ding and Ping Luo	15372
<i>Open Information Extraction via Chunks</i>	
Kuicai Dong, Aixin Sun, Jung-jae Kim and Xiaoli Li	15390
<i>Rethinking Word-Level Auto-Completion in Computer-Aided Translation</i>	
Xingyu Chen, Lemao Liu, Guoping Huang, Zhirui Zhang, Mingming Yang, Shuming Shi and Rui Wang	15405
<i>Automatic Transcription of Handwritten Old Occitan Language</i>	
Esteban Garces Arias, Vallari Pai, Matthias Schöffel, Christian Heumann and Matthias ASSenmacher	15416

<i>CorefPrompt: Prompt-based Event Coreference Resolution by Measuring Event Type and Argument Compatibilities</i>	
Sheng Xu, Peifeng Li and Qiaoming Zhu	15440
<i>Anaphor Assisted Document-Level Relation Extraction</i>	
Chonggang Lu, Richong Zhang, Kai Sun, Jaein Kim, Cunwang Zhang and Yongyi Mao ...	15453
<i>FinEntity: Entity-level Sentiment Classification for Financial Texts</i>	
Yixuan Tang, Yi Yang, Allen H Huang, Andy Tam and Justin Z. Tang	15465
<i>All Things Considered: Detecting Partisan Events from News Media with Cross-Article Comparison</i>	
Yujian Liu, Xinliang Frederick Zhang, Kaijian Zou, Ruihong Huang, Nicholas Beauchamp and Lu Wang	15472
<i>Rationale-Enhanced Language Models are Better Continual Relation Learners</i>	
Weimin Xiong, Yifan Song, Peiyi Wang and Sujian Li	15489
<i>BanglaAbuseMeme: A Dataset for Bengali Abusive Meme Classification</i>	
Mithun Das and Animesh Mukherjee	15498
<i>ScanDL: A Diffusion Model for Generating Synthetic Scanpaths on Texts</i>	
Lena Sophia Bolliger, David Robert Reich, Patrick Haller, Deborah Noemie Jakobi, Paul Prasse and Lena Ann Jäger	15513
<i>From Values to Opinions: Predicting Human Behaviors and Stances Using Value-Injected Large Language Models</i>	
Dongjun Kang, Joonsuk Park, Yohan Jo and JinYeong Bak	15539
<i>Analyzing Film Adaptation through Narrative Alignment</i>	
Tanzir Pial, Shahreen Salim Aunti, Charuta Pethe, Allen Kim and Steven Skiena	15560
<i>Inverse Scaling Can Become U-Shaped</i>	
Jason Wei, Najoung Kim, Yi Tay and Quoc V Le	15580
<i>Nearest Neighbor Machine Translation is Meta-Optimizer on Output Projection Layer</i>	
Ruize Gao, Zhirui Zhang, Yichao Du, Lemao Liu and Rui Wang	15592
<i>Variance Matters: Detecting Semantic Differences without Corpus/Word Alignment</i>	
Ryo Nagata, Hiroya Takamura, Naoki Otani and Yoshifumi Kawasaki	15609
<i>MolCA: Molecular Graph-Language Modeling with Cross-Modal Projector and Uni-Modal Adapter</i>	
Zhiyuan Liu, Sihang Li, Yanchen Luo, Hao Fei, Yixin Cao, Kenji Kawaguchi, Xiang Wang and Tat-Seng Chua	15623
<i>A Training-Free Debiasing Framework with Counterfactual Reasoning for Conversational Emotion Detection</i>	
Geng Tu, Ran Jing, Bin Liang, Min Yang, Kam-Fai Wong and Ruifeng Xu	15639
<i>Self-ICL: Zero-Shot In-Context Learning with Self-Generated Demonstrations</i>	
Wei-Lin Chen, Cheng-Kuang Wu, Yun-Nung Chen and Hsin-Hsi Chen	15651
<i>Learning Knowledge-Enhanced Contextual Language Representations for Domain Natural Language Understanding</i>	
Taolin Zhang, Ruyao Xu, Chengyu Wang, Zhongjie Duan, Cen Chen, Minghui Qiu, Dawei Cheng, Xiaofeng He and Weining Qian	15663
<i>ScdNER: Span-Based Consistency-Aware Document-Level Named Entity Recognition</i>	
Ying Wei and Qi Li	15677

<i>MQuAKE: Assessing Knowledge Editing in Language Models via Multi-Hop Questions</i>	
Zexuan Zhong, Zhengxuan Wu, Christopher D Manning, Christopher Potts and Danqi Chen	15686
<i>Stance Detection on Social Media with Background Knowledge</i>	
Ang Li, Bin Liang, Jingqian Zhao, Bowen Zhang, Min Yang and Ruifeng Xu	15703
<i>Vision-Enhanced Semantic Entity Recognition in Document Images via Visually-Asymmetric Consistency Learning</i>	
Hao Wang, Xiahua Chen, Rui Wang and Chenhui Chu	15718
<i>NormDial: A Comparable Bilingual Synthetic Dialog Dataset for Modeling Social Norm Adherence and Violation</i>	
Oliver Li, Mallika Subramanian, Arkadiy Saakyan, Sky CH-Wang and Smaranda Muresan	15732
<i>ClimateBERT-NetZero: Detecting and Assessing Net Zero and Reduction Targets</i>	
Tobias Schimanski, Julia Bingler, Mathias Kraus, Camilla Hyslop and Markus Leippold	15745
<i>Leap-of-Thought: Accelerating Transformers via Dynamic Token Routing</i>	
Yeachen Kim, Junho Kim, Jun-Hyung Park, Mingyu Lee and SangKeun Lee	15757
<i>Reinforcement Replaces Supervision: Query focused Summarization using Deep Reinforcement Learning</i>	
Swaroop Nath, Pushpak Bhattacharyya and Harshad Khadilkar	15770
<i>Fair Text Classification with Wasserstein Independence</i>	
Thibaud Leteno, Antoine Gourru, Charlotte Laclau, Rémi Emonet and Christophe Gravier	15790
<i>TacoPrompt: A Collaborative Multi-Task Prompt Learning Method for Self-Supervised Taxonomy Completion</i>	
Hongyuan Xu, Ciyi Liu, Yuhang Niu, Yunong Chen, Xiangrui Cai, Yanlong Wen and Xiaojie Yuan	15804
<i>An Attribution Method for Siamese Encoders</i>	
Lucas Moeller, Dmitry Nikolaev and Sebastian Padó	15818
<i>Global Voices, Local Biases: Socio-Cultural Prejudices across Languages</i>	
Anjishnu Mukherjee, Chahat Raj, Ziwei Zhu and Antonios Anastasopoulos	15828
<i>Graph vs. Sequence: An Empirical Study on Knowledge Forms for Knowledge-Grounded Dialogue</i>	
Yizhe Yang, Heyan Huang, Yuhang Liu and Yang Gao	15846
<i>Are Compressed Language Models Less Subgroup Robust?</i>	
Leonidas Gee, Andrea Zugarini and Novi Quadrianto	15859
<i>Length Does Matter: Summary Length can Bias Summarization Metrics</i>	
Xiaobo Guo and Soroush Vosoughi	15869
<i>NL2TL: Transforming Natural Languages to Temporal Logics using Large Language Models</i>	
Yongchao Chen, Rujul Gandhi, Yang Zhang and Chuchu Fan	15880
<i>Reformulating NLP tasks to Capture Longitudinal Manifestation of Language Disorders in People with Dementia.</i>	
Dimitris Gkoumas, Matthew Purver and Maria Liakata	15904
<i>Elevating Code-mixed Text Handling through Auditory Information of Words</i>	
Mamta Mamta, Zishan Ahmad and Asif Ekbal	15918

<i>Predict and Use: Harnessing Predicted Gaze to Improve Multimodal Sarcasm Detection</i>	
Divyank Pratap Tiwari, Diptesh Kanojia, Anupama Ray, Apoorva Nunna and Pushpak Bhattacharyya	15933
<i>Fine-grained Medical Vision-Language Representation Learning for Radiology Report Generation</i>	
Si yuan Wang, Bo Peng, Yichao Liu and Qi Peng	15949
<i>ViT-TTS: Visual Text-to-Speech with Scalable Diffusion Transformer</i>	
Huadai Liu, Rongjie Huang, Xuan Lin, Wenqiang Xu, Maozong Zheng, Hong Chen, Jinzheng He and Zhou Zhao	15957
<i>Consistency Analysis of ChatGPT</i>	
Myeongjun Erik Jang and Thomas Lukasiewicz	15970
<i>Do Differences in Values Influence Disagreements in Online Discussions?</i>	
Michiel van der Meer, Piek Vossen, Catholijn M Jonker and Pradeep Kumar Murukannaiah	15986
<i>Automated Fact-Checking in Dialogue: Are Specialized Models Needed?</i>	
Eric Chamoun, Marzieh Saeidi and Andreas Vlachos	16009
<i>A Digital Language Coherence Marker for Monitoring Dementia</i>	
Dimitris Gkoumas, Adam Tsakalidis and Maria Liakata	16021
<i>Detecting Spoilers in Movie Reviews with External Movie Knowledge and User Networks</i>	
Heng Wang, Wenqian Zhang, Yuyang Bai, Zhaoxuan Tan, Shangbin Feng, Qinghua Zheng and Minnan Luo	16035
<i>Joyful: Joint Modality Fusion and Graph Contrastive Learning for Multimodal Emotion Recognition</i>	
Dongyuan Li, Yusong Wang, Kotaro Funakoshi and Manabu Okumura	16051
<i>HyperRank: Hyperbolic Ranking Model for Unsupervised Keyphrase Extraction</i>	
Mingyang Song, Huafeng Liu and Liping Jing	16070
<i>Assessing the influence of attractor-verb distance on grammatical agreement in humans and language models</i>	
Christos Nikolaos Zacharopoulos, Théo Desbordes and Mathias Sablé-Meyer	16081
<i>Federated Meta-Learning for Emotion and Sentiment Aware Multi-modal Complaint Identification</i>	
Apoorva Singh, Siddarth Chandrasekar, Sriparna Saha and Tanmay Sen	16091
<i>Semantic Similarity Models for Depression Severity Estimation</i>	
Anxo Pérez, Neha Warikoo, Kexin Wang, Javier Parapar and Iryna Gurevych	16104
<i>Hop, Union, Generate: Explainable Multi-hop Reasoning without Rationale Supervision</i>	
Wenting Zhao, Justin T Chiu, Claire Cardie and Alexander M Rush	16119
<i>To Split or Not to Split: Composing Compounds in Contextual Vector Spaces</i>	
Christopher William Jenkins, Filip Miletic and Sabine Schulte im Walde	16131
<i>ToolWriter: Question Specific Tool Synthesis for Tabular Data</i>	
Carlos Gemmell and Jeff Dalton	16137
<i>Interactive Text-to-SQL Generation via Editable Step-by-Step Explanations</i>	
Yuan Tian, Zheng Zhang, Zheng Ning, Toby Jia-Jun Li, Jonathan K. Kummerfeld and Tianyi Zhang	16149

<i>CoCo: Coherence-Enhanced Machine-Generated Text Detection Under Low Resource With Contrastive Learning</i>	
Xiaoming Liu, Zhaohan Zhang, Yichen Wang, Hang Pu, Yu Lan and Chao Shen	16167
<i>AnyTOD: A Programmable Task-Oriented Dialog System</i>	
Jeffrey Zhao, Yuan Cao, Raghav Gupta, Harrison Lee, Abhinav Rastogi, Mingqiu Wang, Hagen Soltau, Izhak Shafran and Yonghui Wu	16189
<i>Can LMs Generalize to Future Data? An Empirical Analysis on Text Summarization</i>	
Chi Seng Cheang, Hou Pong Chan, Derek F. Wong, Xuebo Liu, Zhaocong Li, Yanming Sun, Shudong Liu and Lidia S. Chao	16205
<i>Zero-Shot Multi-Label Topic Inference with Sentence Encoders and LLMs</i>	
Souvika Sarkar, Dongji Feng and Shubhra Kanti Karmaker Santu	16218
<i>TaskDiff: A Similarity Metric for Task-Oriented Conversations</i>	
Ankita Bhaumik, Praveen Venkateswaran, Yara Rizk and Vatche Isahagian	16234
<i>Not all Fake News is Written: A Dataset and Analysis of Misleading Video Headlines</i>	
Yoo Yeon Sung, Jordan Lee Boyd-Graber and Naeemul Hassan	16241
<i>Learning From Free-Text Human Feedback – Collect New Datasets Or Extend Existing Ones?</i>	
Dominic Petrak, Nafise Sadat Moosavi, Ye Tian, Nikolai Rozanov and Iryna Gurevych . . .	16259
<i>Euphemistic Abuse – A New Dataset and Classification Experiments for Implicitly Abusive Language</i>	
Michael Wiegand, Jana Kampfmeier, Elisabeth Eder and Josef Ruppenhofer	16280
<i>Exploring Distributional Shifts in Large Language Models for Code Analysis</i>	
Shushan Arakelyan, Rocktim Jyoti Das, Yi Mao and Xiang Ren	16298
<i>ATHENA: Mathematical Reasoning with Thought Expansion</i>	
JB. Kim, Hazel Kim, Joonghyuk Hahn and Yo-Sub Han	16315
<i>A Benchmark for Reasoning with Spatial Prepositions</i>	
Iulia Maria Comsa and Srini Narayanan	16328
<i>TIMELINE: Exhaustive Annotation of Temporal Relations Supporting the Automatic Ordering of Events in News Articles</i>	
Sarah Alsayyahi and Riza Batista-Navarro	16336
<i>Mitigating Over-Generation for Unsupervised Keyphrase Extraction with Heterogeneous Centrality Detection</i>	
Mingyang Song, Pengyu Xu, Yi Feng, Huafeng Liu and Liping Jing	16349
<i>Towards Interpretable and Efficient Automatic Reference-Based Summarization Evaluation</i>	
Yixin Liu, Alexander Fabbri, Yilun Zhao, Pengfei Liu, Shafiq Joty, Chien-Sheng Wu, Caiming Xiong and Dragomir Radev	16360
<i>MAUD: An Expert-Annotated Legal NLP Dataset for Merger Agreement Understanding</i>	
Steven H Wang, Antoine Scardigli, Leonard Tang, Wei Chen, Dmitry Levkin, Anya Chen, Spencer Ball, Thomas Woodside, Oliver Zhang and Dan Hendrycks	16369
<i>PK-ICR: Persona-Knowledge Interactive Multi-Context Retrieval for Grounded Dialogue</i>	
Minsik Oh, Joosung Lee, Jiwei Li and Guoyin Wang	16383
<i>More Than Spoken Words: Nonverbal Message Extraction and Generation</i>	
Dian Yu, Xiaoyang Wang, Wanshun Chen, Nan Du, Longyue Wang, Haitao Mi and Dong Yu	16396

<i>Can language models learn analogical reasoning? Investigating training objectives and comparisons to human performance</i>	
Molly Petersen and Lonneke van der Plas	16414
<i>FAME: Flexible, Scalable Analogy Mappings Engine</i>	
Shahar Jacob, Chen Shani and Dafna Shahaf	16426
<i>A Self-training Framework for Automated Medical Report Generation</i>	
Siyuan Wang, Zheng Liu and Bo Peng	16443
<i>A Picture is Worth a Thousand Words: Language Models Plan from Pixels</i>	
Anthony Zhe Liu, Lajanugen Logeswaran, Sungryull Sohn and Honglak Lee	16450
<i>Interpreting and Exploiting Functional Specialization in Multi-Head Attention under Multi-task Learning</i>	
Chong Li, Shaonan Wang, Yunhao Zhang, Jiajun Zhang and Chengqing Zong	16460
<i>Multilingual Previously Fact-Checked Claim Retrieval</i>	
Matúš Pikuliak, Ivan Srba, Robert Moro, Timo Hromadka, Timotej Smoleň, Martin Melišek, Ivan Vykopal, Jakub Simko, Juraj Podroužek and Maria Bielikova	16477
<i>ALCAP: Alignment-Augmented Music Captioner</i>	
Zihao He, Weituo Hao, Wei-Tsung Lu, Changyou Chen, Kristina Lerman and Xuchen Song	16501
<i>Do Transformers Parse while Predicting the Masked Word?</i>	
Haoyu Zhao, Abhishek Panigrahi, Rong Ge and Sanjeev Arora	16513
<i>Composable Text Controls in Latent Space with ODEs</i>	
Guangyi Liu, Zeyu Feng, Yuan Gao, Zichao Yang, Xiaodan Liang, Junwei Bao, Xiaodong He, Shuguang Cui, Zhen Li and Zhiting Hu	16543
<i>P5: Plug-and-Play Persona Prompting for Personalized Response Selection</i>	
Joosung Lee, Minsik Oh and Donghun Lee	16571
<i>Reader: Model-based language-instructed reinforcement learning</i>	
Nicola Dainese, Pekka Marttinen and Alexander Ilin	16583
<i>Adapting Offline Speech Translation Models for Streaming with Future-Aware Distillation and Inference</i>	
Biao Fu, Minpeng Liao, Kai Fan, Zhongqiang Huang, Boxing Chen, Yidong Chen and Xiaodong Shi	16600
<i>Relation-aware Ensemble Learning for Knowledge Graph Embedding</i>	
Ling Yue, Yongqi Zhang, Quanming Yao, Yong Li, Xian Wu, Ziheng Zhang, Zhenxi Lin and Yefeng Zheng	16620
<i>GenEx: A Commonsense-aware Unified Generative Framework for Explainable Cyberbullying Detection</i>	
Krishanu Maity, Raghav Jain, Prince Jha, Sriparna Saha and Pushpak Bhattacharyya	16632
<i>Document-Level Machine Translation with Large Language Models</i>	
Longyue Wang, Chenyang Lyu, Tianbo Ji, Zhirui Zhang, Dian Yu, Shuming Shi and Zhaopeng Tu	16646
<i>Multilingual Simplification of Medical Texts</i>	
Sebastian Antony Joseph, Kathryn Kazanas, Keziah Reina, Vishnesh J Ramanathan, Wei Xu, Byron C Wallace and Junyi Jessy Li	16662

<i>When Reviewers Lock Horns: Finding Disagreements in Scientific Peer Reviews</i>	
Sandeep Kumar, Tirthankar Ghosal and Asif Ekbal	16693
<i>Argue with Me Tersely: Towards Sentence-Level Counter-Argument Generation</i>	
Jiayu Lin, Rong Ye, Meng Han, Qi Zhang, Ruofei Lai, Xinyu Zhang, Zhao Cao, Xuanjing Huang and zhongyu wei	16705
<i>JASMINE: Arabic GPT Models for Few-Shot Learning</i>	
El Moatez Billah Nagoudi, Muhammad Abdul-Mageed, AbdelRahim A. Elmadany, Alcides Alcoba Inciarte and Md Tawkat Islam Khondaker	16721
<i>NLI4CT: Multi-Evidence Natural Language Inference for Clinical Trial Reports</i>	
Mael Jullien, Marco Valentino, Hannah Ruth Frost, Paul O'Regan, Dónal Landers and Andre Freitas	16745
<i>Addressing Linguistic Bias through a Contrastive Analysis of Academic Writing in the NLP Domain</i>	
Robert Ridley, Zhen Wu, Jianbing Zhang, Shujian Huang and Xinyu Dai	16765
<i>RobustGEC: Robust Grammatical Error Correction Against Subtle Context Perturbation</i>	
Yue Zhang, Leyang Cui, Enbo Zhao, Wei Bi and Shuming Shi	16780
<i>Detecting Propaganda Techniques in Code-Switched Social Media Text</i>	
Muhammad Umar Salman, Asif Hanif, Shady Shehata and Preslav Nakov	16794
<i>Speech Recognition and Meaning Interpretation: Towards Disambiguation of Structurally Ambiguous Spoken Utterances in Indonesian</i>	
Ruhiyah Faradishi Widiaputri, Ayu Purwarianti, Dessi Puji Lestari, Kurniawati Azizah, Dipta Tanaya and Sakriani Sakti	16813
<i>Target-Agnostic Gender-Aware Contrastive Learning for Mitigating Bias in Multilingual Machine Translation</i>	
Minwoo Lee, Hyukhun Koh, Kang-il Lee, Dongdong Zhang, Minsung Kim and Kyomin Jung	16825
<i>Code-Switching Metrics Using Intonation Units</i>	
Rebecca Pattichis, Dora LaCasse, Sonya Mitrovich Trawick and Rena Torres Cacoullos . . .	16840