

# **Findings of ACL: ACL-IJCNLP 2021**

The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021)

Online  
1-6 August 2021

Volume 1 of 7

ISBN: 978-1-7138-3696-4

**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© (2021) by the Association for Computational Linguistics  
All rights reserved.

Printed with permission by Curran Associates, Inc. (2021)

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

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

Phone: 1-570-476-8006  
Fax: 1-570-476-0860

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

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

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

## Table of Contents

<i>Explainable Inference Over Grounding-Abstract Chains for Science Questions</i> Mokanarangan Thayaparan, Marco Valentino and André Freitas .....	1
<i>LV-BERT: Exploiting Layer Variety for BERT</i> Weihao Yu, Zihang Jiang, Fei Chen, Qibin Hou and Jiashi Feng .....	13
<i>Few-Shot Event Detection with Prototypical Amortized Conditional Random Field</i> Xin Cong, Shiyao Cui, Bowen Yu, Tingwen Liu, Wang Yubin and Bin Wang .....	28
<i>LUX (Linguistic aspects Under eXamination): Discourse Analysis for Automatic Fake News Classification</i> Lucas Azevedo, Mathieu d’Aquin, Brian Davis and Manel Zarrouk .....	41
<i>Diagnosing Transformers in Task-Oriented Semantic Parsing</i> Shrey Desai and Ahmed Aly .....	57
<i>Semantic Relation-aware Difference Representation Learning for Change Captioning</i> Yunbin Tu, Tingting Yao, Liang Li, Jiedong Lou, Shengxiang Gao, Zhengtao YU and Chenggang Yan .....	63
<i>The Authors Matter: Understanding and Mitigating Implicit Bias in Deep Text Classification</i> Haochen Liu, Wei Jin, Hamid Karimi, Zitao Liu and Jiliang Tang .....	74
<i>From What to Why: Improving Relation Extraction with Rationale Graph</i> Zhenyu Zhang, Bowen Yu, Xiaobo Shu, Xue Mengge, Tingwen Liu and Li Guo .....	86
<i>More Parameters? No Thanks!</i> Zeeshan Khan, Kartheek Akella, Vinay Namboodiri and C V Jawahar .....	96
<i>SyGNS: A Systematic Generalization Testbed Based on Natural Language Semantics</i> Hitomi Yanaka, Koji Mineshima and Kentaro Inui .....	103
<i>Fully Non-autoregressive Neural Machine Translation: Tricks of the Trade</i> Jiatao Gu and Xiang Kong .....	120
<i>Generate, Prune, Select: A Pipeline for Counterspeech Generation against Online Hate Speech</i> Wanzheng Zhu and Suma Bhat .....	134
<i>REPT: Bridging Language Models and Machine Reading Comprehension via Retrieval-Based Pre-training</i> Fangkai Jiao, Yangyang Guo, Yilin Niu, Feng Ji, Feng-Lin Li and Liqiang Nie .....	150
<i>CasEE: A Joint Learning Framework with Cascade Decoding for Overlapping Event Extraction</i> Jiawei Sheng, Shu Guo, Bowen Yu, Qian Li, Yiming Hei, Lihong Wang, Tingwen Liu and Hongbo Xu .....	164
<i>Discovering Topics in Long-tailed Corpora with Causal Intervention</i> Xiaobao Wu, Chunping Li and Yishu Miao .....	175
<i>More than just Frequency? Demasking Unsupervised Hypernymy Prediction Methods</i> Thomas Bott, Dominik Schlechtweg and Sabine Schulte im Walde .....	186
<i>WikiTableT: A Large-Scale Data-to-Text Dataset for Generating Wikipedia Article Sections</i> Mingda Chen, Sam Wiseman and Kevin Gimpel .....	193

<i>CoDesc: A Large Code–Description Parallel Dataset</i>	
Masum Hasan, Tanveer Muttaqueen, Abdullah Al Ishtiaq, Kazi Sajeed Mehrab, Md. Mahim Anjum Haque, Tahmid Hasan, Wasi Ahmad, Anindya Iqbal and Rifat Shahriyar .....	210
<i>Deep Cognitive Reasoning Network for Multi-hop Question Answering over Knowledge Graphs</i>	
Jianyu Cai, Zhanqiu Zhang, Feng Wu and Jie Wang .....	219
<i>GoG: Relation-aware Graph-over-Graph Network for Visual Dialog</i>	
Feilong Chen, Xiuyi Chen, Fandong Meng, Peng Li and Jie Zhou .....	230
<i>Joint Optimization of Tokenization and Downstream Model</i>	
Tatsuya Hiraoka, Sho Takase, Kei Uchiumi, Atsushi Keyaki and Naoaki Okazaki .....	244
<i>How does Attention Affect the Model?</i>	
Cheng Zhang, Qiuchi Li, Lingyu Hua and Dawei Song .....	256
<i>Contrastive Attention for Automatic Chest X-ray Report Generation</i>	
Fenglin Liu, Changchang Yin, Xian Wu, Shen Ge, Ping Zhang and Xu Sun .....	269
<i>O2NA: An Object-Oriented Non-Autoregressive Approach for Controllable Video Captioning</i>	
Fenglin Liu, Xuancheng Ren, Xian Wu, Bang Yang, Shen Ge and Xu Sun .....	281
<i>Better Chinese Sentence Segmentation with Reinforcement Learning</i>	
Srivatsan Srinivasan and Chris Dyer .....	293
<i>Enhancing Transformers with Gradient Boosted Decision Trees for NLI Fine-Tuning</i>	
Benjamin Minixhofer, Milan Gritta and Ignacio Iacobacci .....	303
<i>Empirical Error Modeling Improves Robustness of Noisy Neural Sequence Labeling</i>	
Marcin Namysl, Sven Behnke and Joachim Köhler .....	314
<i>Spatial Dependency Parsing for Semi-Structured Document Information Extraction</i>	
Wonseok Hwang, Jinyeong Yim, Seunghyun Park, Sohee Yang and Minjoon Seo .....	330
<i>Reader-Guided Passage Reranking for Open-Domain Question Answering</i>	
Yuning Mao, Pengcheng He, Xiaodong Liu, Yelong Shen, Jianfeng Gao, Jiawei Han and Weizhu Chen .....	344
<i>Entity-Aware Abstractive Multi-Document Summarization</i>	
Hao Zhou, Weidong Ren, Gongshen Liu, Bo Su and Wei Lu .....	351
<i>LenAtten: An Effective Length Controlling Unit For Text Summarization</i>	
Zhongyi Yu, Zhenghao Wu, Hao Zheng, Zhe XuanYuan, Jefferson Fong and Weifeng Su .....	363
<i>XeroAlign: Zero-shot cross-lingual transformer alignment</i>	
Milan Gritta and Ignacio Iacobacci .....	371
<i>Using Word Embeddings to Analyze Teacher Evaluations: An Application to a Filipino Education Non-Profit Organization</i>	
Francesca Vera .....	382
<i>Relation Classification with Entity Type Restriction</i>	
Shengfei Lyu and huanhuan chen .....	390
<i>Link Prediction on N-ary Relational Facts: A Graph-based Approach</i>	
Quan Wang, Haifeng Wang, Yajuan Lyu and Yong Zhu .....	396

<i>GLGE: A New General Language Generation Evaluation Benchmark</i>	
Dayiheng Liu, Yu Yan, Yeyun Gong, Weizhen Qi, Hang Zhang, Jian Jiao, Weizhu Chen, Jie Fu, Linjun Shou, Ming Gong, Pengcheng Wang, Jiusheng Chen, Daxin Jiang, Jiancheng Lv, Ruofei Zhang, Winnie Wu, Ming Zhou and Nan Duan . . . . .	408
<i>AMBERT: A Pre-trained Language Model with Multi-Grained Tokenization</i>	
Xinsong Zhang, Pengshuai Li and Hang Li . . . . .	421
<i>Multimodal Incremental Transformer with Visual Grounding for Visual Dialogue Generation</i>	
Feilong Chen, Fandong Meng, Xiuyi Chen, Peng Li and Jie Zhou . . . . .	436
<i>Retrieve &amp; Memorize: Dialog Policy Learning with Multi-Action Memory</i>	
YunHao Li, Yunyi Yang, Xiaojun Quan and Jianxing Yu . . . . .	447
<i>Adapt-and-Distill: Developing Small, Fast and Effective Pretrained Language Models for Domains</i>	
Yunzhi Yao, Shaohan Huang, Wenhui Wang, Li Dong and Furu Wei . . . . .	460
<i>Decoupling Adversarial Training for Fair NLP</i>	
Xudong Han, Timothy Baldwin and Trevor Cohn . . . . .	471
<i>GO FIGURE: A Meta Evaluation of Factuality in Summarization</i>	
Saadia Gabriel, Asli Celikyilmaz, Rahul Jha, Yejin Choi and Jianfeng Gao . . . . .	478
<i>DNN-driven Gradual Machine Learning for Aspect-term Sentiment Analysis</i>	
Murtadha AHMED, QUN CHEN, Yanyan Wang, youcef nafa, Zhanhuai li and tianyi duan . . . . .	488
<i>Error Detection in Large-Scale Natural Language Understanding Systems Using Transformer Models</i>	
Rakesh Chada, Pradeep Natarajan, Darshan Fofadiya and Prathap Ramachandra . . . . .	498
<i>OutFlip: Generating Examples for Unknown Intent Detection with Natural Language Attack</i>	
DongHyun Choi, Myeong Cheol Shin, EungGyun Kim and Dong Ryeol Shin . . . . .	504
<i>GeoQA: A Geometric Question Answering Benchmark Towards Multimodal Numerical Reasoning</i>	
Jiaqi Chen, Jianheng Tang, Jinghui Qin, Xiaodan Liang, lingbo liu, Eric Xing and Liang Lin . . . . .	513
<i>SIRE: Separate Intra- and Inter-sentential Reasoning for Document-level Relation Extraction</i>	
Shuang Zeng, Yuting Wu and Baobao Chang . . . . .	524
<i>KGPool: Dynamic Knowledge Graph Context Selection for Relation Extraction</i>	
Abhishek Nadgeri, Anson Bastos, Kuldeep Singh, Isaiah Onando Mulang', Johannes Hoffart, Saeedeh Shekarpour and Vijay Saraswat . . . . .	535
<i>Better Combine Them Together! Integrating Syntactic Constituency and Dependency Representations for Semantic Role Labeling</i>	
Hao Fei, Shengqiong Wu, Yafeng Ren, Fei Li and Donghong Ji . . . . .	549
<i>Keep the Primary, Rewrite the Secondary: A Two-Stage Approach for Paraphrase Generation</i>	
Yixuan Su, David Vandyke, Simon Baker, Yan Wang and Nigel Collier . . . . .	560
<i>Contrastive Fine-tuning Improves Robustness for Neural Rankers</i>	
Xiaofei Ma, Cicero Nogueira dos Santos and Andrew O. Arnold . . . . .	570
<i>Cross-Lingual Transfer in Zero-Shot Cross-Language Entity Linking</i>	
Elliot Schumacher, James Mayfield and Mark Dredze . . . . .	583

<i>TellMeWhy: A Dataset for Answering Why-Questions in Narratives</i>	
Yash Kumar Lal, Nathanael Chambers, Raymond Mooney and Niranjan Balasubramanian . . . .	596
<i>Dialogue in the Wild: Learning from a Deployed Role-Playing Game with Humans and Bots</i>	
Kurt Shuster, Jack Urbanek, Emily Dinan, Arthur Szlam and Jason Weston . . . . .	611
<i>Deep Learning against COVID-19: Respiratory Insufficiency Detection in Brazilian Portuguese Speech</i>	
Edresson Casanova, Lucas Gris, Augusto Camargo, Daniel da Silva, Murilo Gazzola, Ester Sabino, Anna Levin, Arnaldo Candido Jr, Sandra Aluisio and Marcelo Finger . . . . .	625
<i>Benchmarking Robustness of Machine Reading Comprehension Models</i>	
Chenglei Si, Ziqing Yang, Yiming Cui, Wentao Ma, Ting Liu and Shijin Wang . . . . .	634
<i>Improving BERT with Syntax-aware Local Attention</i>	
Zhongli Li, Qingyu Zhou, Chao Li, Ke Xu and Yunbo Cao . . . . .	645
<i>A Dialogue-based Information Extraction System for Medical Insurance Assessment</i>	
Shuang Peng, Mengdi Zhou, Minghui Yang, Haitao Mi, Shaosheng Cao, Zujie Wen, Teng Xu, Hongbin Wang and LEI LIU . . . . .	654
<i>Prediction or Comparison: Toward Interpretable Qualitative Reasoning</i>	
Mucheng Ren, Heyan Huang and Yang Gao . . . . .	664
<i>Boundary Detection with BERT for Span-level Emotion Cause Analysis</i>	
Xiangju Li, Wei Gao, Shi Feng, Yifei Zhang and Daling Wang . . . . .	676
<i>On Commonsense Cues in BERT for Solving Commonsense Tasks</i>	
Leyang Cui, Sijie Cheng, Yu Wu and Yue Zhang . . . . .	683
<i>Weakly Supervised Pre-Training for Multi-Hop Retriever</i>	
Yeon Seonwoo, Sang-Woo Lee, Ji-Hoon Kim, Jung-Woo Ha and Alice Oh . . . . .	694
<i>Meet The Truth: Leverage Objective Facts and Subjective Views for Interpretable Rumor Detection</i>	
Jiawen Li, Shiwen Ni and Hung-Yu Kao . . . . .	705
<i>Read, Listen, and See: Leveraging Multimodal Information Helps Chinese Spell Checking</i>	
Heng-Da Xu, Zhongli Li, Qingyu Zhou, Chao Li, Zizhen Wang, Yunbo Cao, Heyan Huang and Xian-Ling Mao . . . . .	716
<i>TransSum: Translating Aspect and Sentiment Embeddings for Self-Supervised Opinion Summarization</i>	
Ke Wang and Xiaojun Wan . . . . .	729
<i>Hashing based Efficient Inference for Image-Text Matching</i>	
Rong-Cheng Tu, Lei Ji, Huaishao Luo, Botian Shi, Heyan Huang, Nan Duan and Xian-Ling Mao	743
<i>Can the Transformer Learn Nested Recursion with Symbol Masking?</i>	
Jean-Philippe Bernardy, Adam Ek and Vladislav Maravev . . . . .	753
<i>Rationalization through Concepts</i>	
Diego Antognini and Boi Faltings . . . . .	761
<i>Parallel Attention Network with Sequence Matching for Video Grounding</i>	
Hao Zhang, Aixin Sun, Wei Jing, Liangli Zhen, Joey Tianyi Zhou and Siow Mong Rick Goh . .	776

<i>MusicBERT: Symbolic Music Understanding with Large-Scale Pre-Training</i> Mingliang Zeng, Xu Tan, Rui Wang, Zeqian Ju, Tao Qin and Tie-Yan Liu .....	791
<i>Evaluating the Efficacy of Summarization Evaluation across Languages</i> Fajri Koto, Jey Han Lau and Timothy Baldwin .....	801
<i>CoMAE: A Multi-factor Hierarchical Framework for Empathetic Response Generation</i> Chujie Zheng, Yong Liu, Wei Chen, Yongcai Leng and Minlie Huang .....	813
<i>UniKeyphrase: A Unified Extraction and Generation Framework for Keyphrase Prediction</i> Huanqin Wu, Wei Liu, Lei Li, Dan Nie, Tao Chen, Feng Zhang and Di Wang .....	825
<i>As Good as New. How to Successfully Recycle English GPT-2 to Make Models for Other Languages</i> Wietse de Vries and Malvina Nissim .....	836
<i>Can Cognate Prediction Be Modelled as a Low-Resource Machine Translation Task?</i> Clémentine Fourrier, Rachel Bawden and Benoît Sagot .....	847
<i>What if This Modified That? Syntactic Interventions with Counterfactual Embeddings</i> Mycal Tucker, Peng Qian and Roger Levy .....	862
<i>Investigating Text Simplification Evaluation</i> Laura Vásquez-Rodríguez, Matthew Shardlow, Piotr Przybyła and Sophia Ananiadou .....	876
<i>COM2SENSE: A Commonsense Reasoning Benchmark with Complementary Sentences</i> Shikhar Singh, Nuan Wen, Yu Hou, Pegah Alipoormolabashi, Te-lin Wu, Xuezhe Ma and Nanyun Peng .....	883
<i>Towards Knowledge-Grounded Counter Narrative Generation for Hate Speech</i> Yi-Ling Chung, Serra Sinem Tekiroğlu and Marco Guerini .....	899
<i>SOLID: A Large-Scale Semi-Supervised Dataset for Offensive Language Identification</i> Sara Rosenthal, Pepa Atanasova, Georgi Karadzhov, Marcos Zampieri and Preslav Nakov ....	915
<i>RealFormer: Transformer Likes Residual Attention</i> Ruining He, Anirudh Ravula, Bhargav Kanagal and Joshua Ainslie .....	929
<i>Promoting Graph Awareness in Linearized Graph-to-Text Generation</i> Alexander Miserlis Hoyle, Ana Marasović and Noah A. Smith .....	944
<i>Predicting cross-linguistic adjective order with information gain</i> William Dyer, Richard Futrell, Zoey Liu and Greg Scontras .....	957
<i>A Survey of Data Augmentation Approaches for NLP</i> Steven Feng, Varun Gangal, Jason Wei, Sarath Chandar, Soroush Vosoughi, Teruko Mitamura and Eduard Hovy .....	968
<i>Why Machine Reading Comprehension Models Learn Shortcuts?</i> Yuxuan Lai, Chen Zhang, Yansong Feng, Quzhe Huang and Dongyan Zhao .....	989
<i>Handling Cross- and Out-of-Domain Samples in Thai Word Segmentation</i> Peerat Limkonchotiwat, Wannaphong Phatthiyaphaibun, Raheem Sarwar, Ekapol Chuangsuwanich and Sarana Nutanong .....	1003
<i>Sensei: Self-Supervised Sensor Name Segmentation</i> Jiaman Wu, Dezhi Hong, Rajesh Gupta and Jingbo Shang .....	1017

<i>Frustratingly Simple Few-Shot Slot Tagging</i>	Jianqiang Ma, ZEYU YAN, Chang Li and Yang Zhang .....	1028
<i>Medical Code Assignment with Gated Convolution and Note-Code Interaction</i>	Shaoxiong Ji, Shirui Pan and Pekka Marttinen .....	1034
<i>Dynamic Semantic Graph Construction and Reasoning for Explainable Multi-hop Science Question Answering</i>	Weiwen Xu, Huihui Zhang, Deng Cai and Wai Lam .....	1044
<i>Addressing Inquiries about History: An Efficient and Practical Framework for Evaluating Open-domain Chatbot Consistency</i>	Zekang Li, Jinchao Zhang, Zhengcong Fei, Yang Feng and Jie Zhou .....	1057
<i>Investigating the Reordering Capability in CTC-based Non-Autoregressive End-to-End Speech Translation</i>	Shun-Po Chuang, Yung-Sung Chuang, Chih-Chiang Chang and Hung-yi Lee .....	1068
<i>Code Summarization with Structure-induced Transformer</i>	Hongqiu Wu, Hai Zhao and Min Zhang .....	1078
<i>Scheduled Dialog Policy Learning: An Automatic Curriculum Learning Framework for Task-oriented Dialog System</i>	Sihong Liu, Jinchao Zhang, Keqing He, Weiran Xu and Jie Zhou .....	1091
<i>Do Explanations Help Users Detect Errors in Open-Domain QA? An Evaluation of Spoken vs. Visual Explanations</i>	Ana Valeria González, Gagan Bansal, Angela Fan, Yashar Mehdad, Robin Jia and Srinivasan Iyer	1103
<i>OntoEA: Ontology-guided Entity Alignment via Joint Knowledge Graph Embedding</i>	Yuejia Xiang, Ziheng Zhang, Jiaoyan Chen, Xi Chen, Zhenxi Lin and Yefeng Zheng .....	1117
<i>Learning Algebraic Recombination for Compositional Generalization</i>	Chenyao Liu, Shengnan An, Zeqi Lin, Qian Liu, Bei Chen, Jian-Guang LOU, Lijie Wen, Nanning Zheng and Dongmei Zhang .....	1129
<i>Out of Order: How important is the sequential order of words in a sentence in Natural Language Understanding tasks?</i>	Thang Pham, Trung Bui, Long Mai and Anh Nguyen .....	1145
<i>RevCore: Review-Augmented Conversational Recommendation</i>	Yu Lu, Junwei Bao, Yan Song, Zichen Ma, Shuguang Cui, Youzheng Wu and Xiaodong He ..	1161
<i>Awakening Latent Grounding from Pretrained Language Models for Semantic Parsing</i>	Qian Liu, Dejian Yang, Jiahui Zhang, Jiaqi Guo, Bin Zhou and Jian-Guang LOU .....	1174
<i>Enhancing Label Correlation Feedback in Multi-Label Text Classification via Multi-Task Learning</i>	Ximing Zhang, Qian-Wen Zhang, Zhao Yan, Ruifang Liu and Yunbo Cao .....	1190
<i>Fusing Context Into Knowledge Graph for Commonsense Question Answering</i>	Yichong Xu, Chenguang Zhu, Ruochen Xu, Yang Liu, Michael Zeng and Xuedong Huang ...	1201
<i>Unsupervised Energy-based Adversarial Domain Adaptation for Cross-domain Text Classification</i>	Han Zou, Jianfei Yang and Xiaojian Wu .....	1208



<i>Survival text regression for time-to-event prediction in conversations</i>	
Christine De Kock and Andreas Vlachos .....	1219
<i>Unsupervised Knowledge Selection for Dialogue Generation</i>	
Xiuyi Chen, Feilong Chen, Fandong Meng, Peng Li and Jie Zhou .....	1230
<i>Minimax and Neyman–Pearson Meta-Learning for Outlier Languages</i>	
Edoardo Maria Ponti, Rahul Aralikkatte, Disha Shrivastava, Siva Reddy and Anders Søgaard ..	1245
<i>On-the-Fly Attention Modulation for Neural Generation</i>	
Yue Dong, Chandra Bhagavatula, Ximing Lu, Jena D. Hwang, Antoine Bosselut, Jackie Chi Kit Cheung and Yejin Choi .....	1261
<i>Grammar-Constrained Neural Semantic Parsing with LR Parsers</i>	
Artur Baranowski and Nico Hochgeschwender .....	1275
<i>Enhanced Metaphor Detection via Incorporation of External Knowledge Based on Linguistic Theories</i>	
Chang Su, Kechun Wu and Yijiang Chen .....	1280
<i>Controlling Text Edition by Changing Answers of Specific Questions</i>	
Lei Sha, Patrick Hohenecker and Thomas Lukasiewicz .....	1288
<i>Grammar-Based Patches Generation for Automated Program Repair</i>	
Yu Tang, Long Zhou, Ambrosio Blanco, Shujie Liu, Furu Wei, Ming Zhou and Muyun Yang ..	1300
<i>Manual Evaluation Matters: Reviewing Test Protocols of Distantly Supervised Relation Extraction</i>	
Tianyu Gao, Xu Han, Yuzhuo Bai, Keyue Qiu, Zhiyu Xie, Yankai Lin, Zhiyuan Liu, Peng Li, Maosong Sun and Jie Zhou .....	1306
<i>GCRC: A New Challenging MRC Dataset from Gaokao Chinese for Explainable Evaluation</i>	
Hongye Tan, xiaoyue Wang, Yu Ji, Ru Li, Xiaoli Li, Zhiwei Hu, Yunxiao Zhao and Xiaoqi Han	1319
<i>Zero-shot Label-Aware Event Trigger and Argument Classification</i>	
Hongming Zhang, Haoyu Wang and Dan Roth .....	1331
<i>Incorporating Global Information in Local Attention for Knowledge Representation Learning</i>	
Yu Zhao, Han Zhou, Ruobing Xie, Fuzhen Zhuang, Qing Li and Ji Liu .....	1341
<i>Exploiting Position Bias for Robust Aspect Sentiment Classification</i>	
Fang Ma, Chen Zhang and Dawei Song .....	1352
<i>MRN: A Locally and Globally Mention-Based Reasoning Network for Document-Level Relation Extraction</i>	
Jingye Li, Kang Xu, Fei Li, Hao Fei, Yafeng Ren and Donghong Ji .....	1359
<i>Adversary-Aware Rumor Detection</i>	
Yun-Zhu Song, Yi-Syuan Chen, Yi-Ting Chang, Shao-Yu Weng and Hong-Han Shuai .....	1371
<i>LICHEE: Improving Language Model Pre-training with Multi-grained Tokenization</i>	
Weidong Guo, Mingjun Zhao, Lusheng Zhang, Di Niu, Jinwen Luo, Zhenhua Liu, Zhenyang Li and Jianbo Tang .....	1383
<i>Detecting Hallucinated Content in Conditional Neural Sequence Generation</i>	
Chunting Zhou, Graham Neubig, Jiatao Gu, Mona Diab, Francisco Guzmán, Luke Zettlemoyer and Marjan Ghazvininejad .....	1393

<i>K-Adapter: Infusing Knowledge into Pre-Trained Models with Adapters</i>	
Ruize Wang, Duyu Tang, Nan Duan, Zhongyu Wei, Xuanjing Huang, Jianshu Ji, Guihong Cao, Daxin Jiang and Ming Zhou . . . . .	1405
<i>Global Attention Decoder for Chinese Spelling Error Correction</i>	
Zhao Guo, Yuan Ni, Keqiang Wang, Wei Zhu and GUOTONG XIE . . . . .	1419
<i>Jointly Identifying Rhetoric and Implicit Emotions via Multi-Task Learning</i>	
Xin Chen, Zhen Hai, Deyu Li, Suge Wang and Dian Wang . . . . .	1429
<i>Exploring the Role of Context in Utterance-level Emotion, Act and Intent Classification in Conversations: An Empirical Study</i>	
Deepanway Ghosal, Navonil Majumder, Rada Mihalcea and Soujanya Poria . . . . .	1435
<i>Encouraging Neural Machine Translation to Satisfy Terminology Constraints</i>	
Melissa Ailem, Jingshu Liu and Raheel Qader . . . . .	1450
<i>BertGCN: Transductive Text Classification by Combining GNN and BERT</i>	
Yuxiao Lin, Yuxian Meng, Xiaofei Sun, Qinghong Han, Kun Kuang, Jiwei Li and Fei Wu . . .	1456
<i>Putting words into the system’s mouth: A targeted attack on neural machine translation using monolingual data poisoning</i>	
Jun Wang, Chang Xu, Francisco Guzmán, Ahmed El-Kishky, Yuqing Tang, Benjamin Rubinstein and Trevor Cohn . . . . .	1463
<i>Semantic and Syntactic Enhanced Aspect Sentiment Triplet Extraction</i>	
Zhexue Chen, Hong Huang, Bang Liu, Xuanhua Shi and Hai Jin . . . . .	1474
<i>UserAdapter: Few-Shot User Learning in Sentiment Analysis</i>	
Wanjun Zhong, Duyu Tang, Jiahai Wang, Jian Yin and Nan Duan . . . . .	1484
<i>PsyQA: A Chinese Dataset for Generating Long Counseling Text for Mental Health Support</i>	
Hao Sun, Zhenru Lin, Chujie Zheng, Siyang Liu and Minlie Huang . . . . .	1489
<i>RiddleSense: Reasoning about Riddle Questions Featuring Linguistic Creativity and Commonsense Knowledge</i>	
Bill Yuchen Lin, Ziyi Wu, Yichi Yang, Dong-Ho Lee and Xiang Ren . . . . .	1504
<i>Learning to Generate Questions by Learning to Recover Answer-containing Sentences</i>	
Seohyun Back, Akhil Kedia, Sai Chetan Chinthakindi, Haejun Lee and Jaegul Choo . . . . .	1516
<i>Learning Slice-Aware Representations with Mixture of Attention</i>	
Cheng Wang, Sungjin Lee, Sunghyun Park, Han Li, Young-Bum Kim and Ruhi Sarikaya . . .	1530
<i>Making Better Use of Bilingual Information for Cross-Lingual AMR Parsing</i>	
Yitao Cai, Zhe Lin and Xiaojun Wan . . . . .	1537
<i>Pushing Paraphrase Away from Original Sentence: A Multi-Round Paraphrase Generation Approach</i>	
Zhe Lin and Xiaojun Wan . . . . .	1548
<i>Few-shot Knowledge Graph-to-Text Generation with Pretrained Language Models</i>	
Junyi Li, Tianyi Tang, Wayne Xin Zhao, Zhicheng Wei, Nicholas Jing Yuan and Ji-Rong Wen	1558
<i>Better Robustness by More Coverage: Adversarial and Mixup Data Augmentation for Robust Finetuning</i>	
Chenglei Si, Zhengyan Zhang, Fanchao Qi, Zhiyuan Liu, Yasheng Wang, Qun Liu and Maosong Sun . . . . .	1569

<i>NAST: A Non-Autoregressive Generator with Word Alignment for Unsupervised Text Style Transfer</i> Fei Huang, Zikai Chen, Chen Henry Wu, Qihan Guo, Xiaoyan Zhu and Minlie Huang . . . . .	1577
<i>HyKnow: End-to-End Task-Oriented Dialog Modeling with Hybrid Knowledge Management</i> Silin Gao, Ryuichi Takanobu, Wei Peng, Qun Liu and Minlie Huang . . . . .	1591
<i>Target-oriented Fine-tuning for Zero-Resource Named Entity Recognition</i> Ying Zhang, Fandong Meng, Yufeng Chen, Jinan Xu and Jie Zhou . . . . .	1603
<i>BERT-Defense: A Probabilistic Model Based on BERT to Combat Cognitively Inspired Orthographic Adversarial Attacks</i> Yannik Keller, Jan Mackensen and Steffen Eger . . . . .	1616
<i>Event Detection as Graph Parsing</i> Jianye xie, Haotong Sun, Junsheng Zhou, Weiguang Qu and Xinyu Dai . . . . .	1630
<i>Toward Fully Exploiting Heterogeneous Corpus:A Decoupled Named Entity Recognition Model with Two-stage Training</i> Yun Hu, Yeshuang Zhu, Jinchao Zhang, Changwen Zheng and Jie Zhou . . . . .	1641
<i>Discriminative Reasoning for Document-level Relation Extraction</i> Wang Xu, Kehai Chen and Tiejun Zhao . . . . .	1653
<i>Meta-Learning Adversarial Domain Adaptation Network for Few-Shot Text Classification</i> Chengcheng Han, Zeqiu Fan, Dongxiang Zhang, Minghui Qiu, Ming Gao and Aoying Zhou . . . . .	1664
<i>Documents Representation via Generalized Coupled Tensor Chain with the Rotation Group constraint</i> Igor Vorona, Anh-Huy Phan, Alexander Panchenko and Andrzej Cichocki . . . . .	1674
<i>Improving Unsupervised Extractive Summarization with Facet-Aware Modeling</i> Xinnian Liang, Shuangzhi Wu, Mu Li and Zhoujun Li . . . . .	1685
<i>Improving Gradient-based Adversarial Training for Text Classification by Contrastive Learning and Auto-Encoder</i> Yao Qiu, Jinchao Zhang and Jie Zhou . . . . .	1698
<i>Multi-Granularity Contrasting for Cross-Lingual Pre-Training</i> Shicheng Li, Pengcheng Yang, Fuli Luo and Jun Xie . . . . .	1708
<i>A Comparison between Pre-training and Large-scale Back-translation for Neural Machine Translation</i> Dandan Huang, Kun Wang and Yue Zhang . . . . .	1718
<i>Bi-Granularity Contrastive Learning for Post-Training in Few-Shot Scene</i> Ruikun Luo, Guanhuan Huang and Xiaojun Quan . . . . .	1733
<i>Fusing Label Embedding into BERT: An Efficient Improvement for Text Classification</i> Yijin Xiong, Yukun Feng, Hao Wu, Hidetaka Kamigaito and Manabu Okumura . . . . .	1743
<i>KACC: A Multi-task Benchmark for Knowledge Abstraction, Concretization and Completion</i> Jie Zhou, Shengding Hu, Xin Lv, Cheng Yang, Zhiyuan Liu, Wei Xu, Jie Jiang, Juanzi Li and Maosong Sun . . . . .	1751
<i>A Query-Driven Topic Model</i> Zheng Fang, Yulan He and Rob Procter . . . . .	1764

<i>How Reliable are Model Diagnostics?</i>	
Vamsi Aribandi, Yi Tay and Donald Metzler .....	1778
<i>Gaussian Process based Deep Dyna-Q approach for Dialogue Policy Learning</i>	
Guanlin Wu, Wenqi Fang, Ji Wang, Jiang Cao, Weidong Bao, Yang Ping, Xiaomin Zhu and Zheng Wang .....	1786
<i>CiteWorth: Cite-Worthiness Detection for Improved Scientific Document Understanding</i>	
Dustin Wright and Isabelle Augenstein .....	1796
<i>Cross-Lingual Cross-Domain Nested Named Entity Evaluation on English Web Texts</i>	
Barbara Plank .....	1808
<i>Counter-Argument Generation by Attacking Weak Premises</i>	
Milad Alshomary, Shahbaz Syed, Arkajit Dhar, Martin Potthast and Henning Wachsmuth . . .	1816
<i>Alternated Training with Synthetic and Authentic Data for Neural Machine Translation</i>	
Rui Jiao, Zonghan Yang, Maosong Sun and Yang Liu .....	1828
<i>Template-Based Named Entity Recognition Using BART</i>	
Leyang Cui, Yu Wu, Jian Liu, Sen Yang and Yue Zhang .....	1835
<i>"Does it Matter When I Think You Are Lying?" Improving Deception Detection by Integrating Interlocutor's Judgements in Conversations</i>	
Huang-Cheng Chou, Woan-Shiuan Chien, Da-Cheng Juan and Chi-Chun Lee .....	1846
<i>High-Quality Dialogue Diversification by Intermittent Short Extension Ensembles</i>	
Zhiwen Tang, Hrishikesh Kulkarni and Grace Hui Yang .....	1861
<i>Structured Refinement for Sequential Labeling</i>	
Yiran Wang, Hiroyuki Shindo, Yuji Matsumoto and Taro Watanabe .....	1873
<i>End-to-End Construction of NLP Knowledge Graph</i>	
Ishani Mondal, Yufang Hou and Charles Jochim .....	1885
<i>Deciphering Implicit Hate: Evaluating Automated Detection Algorithms for Multimodal Hate</i>	
Austin Botelho, Scott Hale and Bertie Vidgen .....	1896
<i>Studying the Evolution of Scientific Topics and their Relationships</i>	
Ana Sabina Uban, Cornelia Caragea and Liviu P. Dinu .....	1908
<i>End-to-End Self-Debiasing Framework for Robust NLU Training</i>	
Abbas Ghaddar, Phillippe Langlais, Mehdi Rezagholizadeh and Ahmad Rashid .....	1923
<i>A Mixed-Method Design Approach for Empirically Based Selection of Unbiased Data Annotators</i>	
Gautam Thakur, Janna Caspersen, Drahomira Herrmannova, Bryan Eaton and Jordan Burdette	1930
<i>An Evaluation of Disentangled Representation Learning for Texts</i>	
Krishnapriya Vishnubhotla, Graeme Hirst and Frank Rudzicz .....	1939
<i>Injecting Knowledge Base Information into End-to-End Joint Entity and Relation Extraction and Coreference Resolution</i>	
Severine Verlinden, Klim Zaporozhets, Johannes Deleu, Thomas Demeester and Chris Develder	1952
<i>Knowing More About Questions Can Help: Improving Calibration in Question Answering</i>	
Shujian Zhang, Chengyue Gong and Eunsol Choi .....	1958

<i>Enhancing Metaphor Detection by Gloss-based Interpretations</i>	
Hai Wan, Jinxia Lin, Jianfeng Du, Dawei Shen and Manrong Zhang .....	1971
<i>Evaluating Word Embeddings with Categorical Modularity</i>	
Sílvia Casacuberta, Karina Halevy and Damián Blasi .....	1982
<i>Attention-based Contextual Language Model Adaptation for Speech Recognition</i>	
Richard Diehl Martinez, Scott Novotney, Ivan Bulyko, Ariya Rastrow, Andreas Stolcke and Ankur Gandhe .....	1994
<i>Annotation and Evaluation of Coreference Resolution in Screenplays</i>	
Sabyasachee Baruah, Sandeep Nallan Chakravarthula and Shrikanth Narayanan .....	2004
<i>Exploring Cross-Lingual Transfer Learning with Unsupervised Machine Translation</i>	
Chao Wang, Judith Gaspers, Thi Ngoc Quynh Do and Hui Jiang .....	2011
<i>Pipeline Signed Japanese Translation Focusing on a Post-positional Particle Complement and Conjugation in a Low-resource Setting</i>	
Ken Yano and Akira Utsumi .....	2021
<i>Language-Mediated, Object-Centric Representation Learning</i>	
Ruocheng Wang, Jiayuan Mao, Samuel Gershman and Jiajun Wu .....	2033
<i>Entheos: A Multimodal Dataset for Studying Enthusiasm</i>	
Carla Viegas and Malihe Alikhani .....	2047
<i>Are Rotten Apples Edible? Challenging Commonsense Inference Ability with Exceptions</i>	
Nam Do and Ellie Pavlick .....	2061
<i>GRICE: A Grammar-based Dataset for Recovering Implicature and Conversational Reasoning</i>	
Zilong Zheng, Shuwen Qiu, Lifeng Fan, Yixin Zhu and Song-Chun Zhu .....	2074
<i>RetroGAN: A Cyclic Post-Specialization System for Improving Out-of-Knowledge and Rare Word Representations</i>	
Pedro Colon-Hernandez, Yida Xin, Henry Lieberman, Catherine Havasi, Cynthia Breazeal and Peter Chin .....	2086
<i>Fusion: Towards Automated ICD Coding via Feature Compression</i>	
Junyu Luo, Cao Xiao, Lucas Glass, Jimeng Sun and Fenglong Ma .....	2096
<i>Automatic Document Sketching: Generating Drafts from Analogous Texts</i>	
Zeiqu Wu, Michel Galley, Chris Brockett, Yizhe Zhang and Bill Dolan .....	2102
<i>Trade the Event: Corporate Events Detection for News-Based Event-Driven Trading</i>	
Zhihan Zhou, Liqian Ma and Han Liu .....	2114
<i>Language-based General Action Template for Reinforcement Learning Agents</i>	
Ryosuke Kohita, Akifumi Wachi, Daiki Kimura, Subhajt Chaudhury, Michiaki Tatsubori and Asim Munawar .....	2125
<i>MiniLMv2: Multi-Head Self-Attention Relation Distillation for Compressing Pretrained Transformers</i>	
Wenhui Wang, Hangbo Bao, Shaohan Huang, Li Dong and Furu Wei .....	2140
<i>Attending via both Fine-tuning and Compressing</i>	
Jie Zhou, Yuanbin Wu, Qin Chen, Xuanjing Huang and liang he .....	2152

<i>Improving Event Causality Identification via Self-Supervised Representation Learning on External Causal Statement</i>	
Xinyu Zuo, Pengfei Cao, Yubo Chen, Kang Liu, Jun Zhao, Weihua Peng and Yuguang Chen .	2162
<i>PAIR: Leveraging Passage-Centric Similarity Relation for Improving Dense Passage Retrieval</i>	
Ruiyang Ren, Shangwen Lv, Yingqi Qu, Jing Liu, Wayne Xin Zhao, QiaoQiao She, Hua Wu, Haifeng Wang and Ji-Rong Wen . . . . .	2173
<i>Is Human Scoring the Best Criteria for Summary Evaluation?</i>	
Oleg Vasilyev and John Bohannon . . . . .	2184
<i>Assessing Dialogue Systems with Distribution Distances</i>	
Jiannan Xiang, Yahui Liu, Deng Cai, Huayang Li, Defu Lian and Lemao Liu . . . . .	2192
<i>Neural Combinatory Constituency Parsing</i>	
Zhousi Chen, Longtu Zhang, Aizhan Imankulova and Mamoru Komachi . . . . .	2199
<i>Learning Shared Semantic Space for Speech-to-Text Translation</i>	
Chi Han, Mingxuan Wang, Heng Ji and Lei Li . . . . .	2214
<i>Empowering Language Understanding with Counterfactual Reasoning</i>	
Fuli Feng, Jizhi Zhang, Xiangnan He, Hanwang Zhang and Tat-Seng Chua . . . . .	2226
<i>Knowledge-Empowered Representation Learning for Chinese Medical Reading Comprehension: Task, Model and Resources</i>	
Taolin Zhang, Chengyu Wang, Minghui Qiu, Bite Yang, Zerui Cai, XIAOFENG HE and jun huang	2237
<i>Correcting Chinese Spelling Errors with Phonetic Pre-training</i>	
Ruiqing Zhang, Chao Pang, Chuanqiang Zhang, Shuohuan Wang, Zhongjun He, Yu Sun, Hua Wu and Haifeng Wang . . . . .	2250
<i>Multi-Lingual Question Generation with Language Agnostic Language Model</i>	
Bingning Wang, Ting Yao, Weipeng Chen, jingfang xu and Xiaochuan Wang . . . . .	2262
<i>Structure-Aware Pre-Training for Table-to-Text Generation</i>	
Xinyu Xing and Xiaojun Wan . . . . .	2273
<i>On the Interplay Between Fine-tuning and Composition in Transformers</i>	
Lang Yu and Allyson Ettinger . . . . .	2279
<i>Lifelong Learning of Topics and Domain-Specific Word Embeddings</i>	
Xiaorui Qin, Yuyin Lu, Yufu Chen and Yanghui Rao . . . . .	2294
<i>Leveraging Argumentation Knowledge Graph for Interactive Argument Pair Identification</i>	
Jian Yuan, Zhongyu Wei, Donghua Zhao, Qi Zhang and Changjian Jiang . . . . .	2310
<i>A Multi-Task Learning Framework for Multi-Target Stance Detection</i>	
Yingjie Li and Cornelia Caragea . . . . .	2320
<i>Confidence-Aware Scheduled Sampling for Neural Machine Translation</i>	
Yijin Liu, Fandong Meng, Yufeng Chen, Jinan Xu and Jie Zhou . . . . .	2327
<i>MA-BERT: Learning Representation by Incorporating Multi-Attribute Knowledge in Transformers</i>	
You Zhang, Jin Wang, Liang-Chih Yu and Xuejie Zhang . . . . .	2338

<i>A Closer Look into the Robustness of Neural Dependency Parsers Using Better Adversarial Examples</i> Yuxuan Wang, Wanxiang Che, Ivan Titov, Shay B. Cohen, Zhilin Lei and Ting Liu . . . . .	2344
<i>P-Stance: A Large Dataset for Stance Detection in Political Domain</i> Yingjie Li, Tiberiu Sosea, Aditya Sawant, Ajith Jayaraman Nair, Diana Inkpen and Cornelia Caragea	2355
<i>WIND: Weighting Instances Differentially for Model-Agnostic Domain Adaptation</i> Xiang Chen, Yue Cao and Xiaojun Wan . . . . .	2366
<i>DocOIE: A Document-level Context-Aware Dataset for OpenIE</i> Kuicai Dong, Zhao Yilin, Aixin Sun, Jung-Jae Kim and Xiaoli Li . . . . .	2377
<i>Event Extraction from Historical Texts: A New Dataset for Black Rebellions</i> Viet Lai, Minh Van Nguyen, Heidi Kaufman and Thien Huu Nguyen . . . . .	2390
<i>Zero-shot Medical Entity Retrieval without Annotation: Learning From Rich Knowledge Graph Semantics</i> Luyang Kong, Christopher Winestock and Parminder Bhatia . . . . .	2401
<i>CONDA: a CONTEXTual Dual-Annotated dataset for in-game toxicity understanding and detection</i> Henry Weld, Guanghao Huang, Jean Lee, Tongshu Zhang, Kunze Wang, Xinghong Guo, Siqu Long, Josiah Poon and Caren Han . . . . .	2406
<i>Adaptive Knowledge-Enhanced Bayesian Meta-Learning for Few-shot Event Detection</i> Shirong Shen, Tongtong Wu, Guilin Qi, Yuan-Fang Li, Gholamreza Haffari and Sheng Bi . . .	2417
<i>Stylized Story Generation with Style-Guided Planning</i> Xiangzhe Kong, Jialiang Huang, Ziquan Tung, Jian Guan and Minlie Huang . . . . .	2430
<i>Dynamic Connected Networks for Chinese Spelling Check</i> Baixin Wang, Wanxiang Che, dayong wu, Shijin Wang, Guoping Hu and Ting Liu . . . . .	2437
<i>A Multi-Level Attention Model for Evidence-Based Fact Checking</i> Canasai Kruengkrai, Junichi Yamagishi and Xin Wang . . . . .	2447
<i>RealTrans: End-to-End Simultaneous Speech Translation with Convolutional Weighted-Shrinking Transformer</i> Kingshan Zeng, Liangyou Li and Qun Liu . . . . .	2461
<i>Training ELECTRA Augmented with Multi-word Selection</i> Jiaming Shen, Jialu Liu, Tianqi Liu, Cong Yu and Jiawei Han . . . . .	2475
<i>REAM<sup>‡</sup>: An Enhancement Approach to Reference-based Evaluation Metrics for Open-domain Dialog Generation</i> Jun Gao, Wei Bi, Ruifeng Xu and Shuming Shi . . . . .	2487
<i>Relation Extraction with Type-aware Map Memories of Word Dependencies</i> Guimin Chen, Yuanhe Tian, Yan Song and Xiang Wan . . . . .	2501
<i>PLATO-2: Towards Building an Open-Domain Chatbot via Curriculum Learning</i> Siqu Bao, Huang He, Fan Wang, Hua Wu, Haifeng Wang, Wenquan Wu, Zhen Guo, Zhibin Liu and Xinchao Xu . . . . .	2513

<i>JointGT: Graph-Text Joint Representation Learning for Text Generation from Knowledge Graphs</i> Pei Ke, Haozhe Ji, Yu Ran, Xin Cui, Liwei Wang, Linfeng Song, Xiaoyan Zhu and Minlie Huang	2526
<i>AdaST: Dynamically Adapting Encoder States in the Decoder for End-to-End Speech-to-Text Translation</i> Wuwei Huang, Dexin Wang and Deyi Xiong	2539
<i>OKGIT: Open Knowledge Graph Link Prediction with Implicit Types</i> . Chandrahas and Partha Talukdar	2546
<i>Multimodal Fusion with Co-Attention Networks for Fake News Detection</i> Yang Wu, Pengwei Zhan, Yunjian Zhang, Liming Wang and Zhen Xu	2560
<i>Joint Multi-Decoder Framework with Hierarchical Pointer Network for Frame Semantic Parsing</i> Xudong Chen, Ce Zheng and Baobao Chang	2570
<i>H-FND: Hierarchical False-Negative Denoising for Distant Supervision Relation Extraction</i> JHIH-WEI CHEN, Tsu-Jui Fu, Chen-Kang Lee and Wei-Yun Ma	2579
<i>GEM: A General Evaluation Benchmark for Multimodal Tasks</i> Lin Su, Nan Duan, Edward Cui, Lei Ji, Chenfei Wu, Huaishao Luo, Yongfei Liu, Ming Zhong, Taroon Bharti and Arun Sacheti	2594
<i>Graph Relational Topic Model with Higher-order Graph Attention Auto-encoders</i> Qianqian Xie, Jimin Huang, Pan Du and Min Peng	2604
<i>Paths to Relation Extraction through Semantic Structure</i> Jonathan Yellin and Omri Abend	2614
<i>Dynamic and Multi-Channel Graph Convolutional Networks for Aspect-Based Sentiment Analysis</i> Shiguan Pang, Yun Xue, Zehao Yan, Weihao Huang and Jinhui Feng	2627
<i>Automatic Text Simplification for Social Good: Progress and Challenges</i> Sanja Stajner	2637
<i>A Neural Edge-Editing Approach for Document-Level Relation Graph Extraction</i> Kohei Makino, Makoto Miwa and Yutaka Sasaki	2653
<i>Dialogue-oriented Pre-training</i> Yi Xu and Hai Zhao	2663
<i>GrantRel: Grant Information Extraction via Joint Entity and Relation Extraction</i> Junyi Bian, Li Huang, Xiaodi Huang, Hong Zhou and Shanfeng Zhu	2674
<i>Enhancing Language Generation with Effective Checkpoints of Pre-trained Language Model</i> Jeonghyeok Park and Hai Zhao	2686
<i>Making Flexible Use of Subtasks: A Multiplex Interaction Network for Unified Aspect-based Sentiment Analysis</i> Guoxin Yu, Xiang Ao, Ling Luo, Min Yang, Xiaofei Sun, Jiwei Li and Qing He	2695
<i>Continual Mixed-Language Pre-Training for Extremely Low-Resource Neural Machine Translation</i> Zihan Liu, Genta Indra Winata and Pascale Fung	2706
<i>Transformer-Exclusive Cross-Modal Representation for Vision and Language</i> Andrew Shin and Takuya Narihira	2719



<i>Two Parents, One Child: Dual Transfer for Low-Resource Neural Machine Translation</i> Meng Zhang, Liangyou Li and Qun Liu .....	2726
<i>Contrastive Aligned Joint Learning for Multilingual Summarization</i> Danqing Wang, Jiaze Chen, Hao Zhou, Xipeng Qiu and Lei Li .....	2739
<i>When Time Makes Sense: A Historically-Aware Approach to Targeted Sense Disambiguation</i> Kaspar Beelen, Federico Nanni, Mariona Coll Ardanuy, Kasma Hosseini, Giorgia Tolfo and Barbara McGillivray .....	2751
<i>Understanding Feature Focus in Multitask Settings for Lexico-semantic Relation Identification</i> Houssam Akhmouch, Gaël Dias and Jose G. Moreno .....	2762
<i>Don't Miss the Labels: Label-semantic Augmented Meta-Learner for Few-Shot Text Classification</i> Qiaoyang Luo, Lingqiao Liu, Yuhao Lin and Wei Zhang .....	2773
<i>Detecting Harmful Memes and Their Targets</i> Shraman Pramanick, Dimitar Dimitrov, Rituparna Mukherjee, Shivam Sharma, Md. Shad Akhtar, Preslav Nakov and Tanmoy Chakraborty .....	2783
<i>Progressive Multi-Granularity Training for Non-Autoregressive Translation</i> Liang Ding, Longyue Wang, Xuebo Liu, Derek F. Wong, Dacheng Tao and Zhaopeng Tu .....	2797
<i>ZmBART: An Unsupervised Cross-lingual Transfer Framework for Language Generation</i> Kaushal Kumar Maurya, Maunendra Sankar Desarkar, Yoshinobu Kano and Kumari Deepshikha	2804
<i>HacRED: A Large-Scale Relation Extraction Dataset Toward Hard Cases in Practical Applications</i> Qiao Cheng, Juntao Liu, Xiaoye Qu, Jin Zhao, Jiaqing Liang, Zhefeng Wang, baoxing Huai, Nicholas Jing Yuan and Yanghua Xiao .....	2819
<i>Do Multilingual Neural Machine Translation Models Contain Language Pair Specific Attention Heads?</i> Zae Myung Kim, Laurent Besacier, Vassilina Nikoulina and Didier Schwab .....	2832
<i>Learning Sequential and Structural Information for Source Code Summarization</i> YunSeok Choi, JinYeong Bak, CheolWon Na and Jee-Hyong Lee .....	2842
<i>Energy-based Unknown Intent Detection with Data Manipulation</i> Yawen Ouyang, Jiasheng Ye, Yu Chen, Xinyu Dai, Shujian Huang and Jiajun CHEN .....	2852
<i>Automatic Rephrasing of Transcripts-based Action Items</i> Amir Cohen, Amir Kantor, Sagi Hilleli and Eyal Kolman .....	2862
<i>MergeDistill: Merging Language Models using Pre-trained Distillation</i> Simran Khanuja, Melvin Johnson and Partha Talukdar .....	2874
<i>On Sparsifying Encoder Outputs in Sequence-to-Sequence Models</i> Biao Zhang, Ivan Titov and Rico Sennrich .....	2888
<i>FrameNet-assisted Noun Compound Interpretation</i> Girishkumar Ponkiya, Diptesh Kanojia, Pushpak Bhattacharyya and Girish Palshikar .....	2901
<i>Hypernym Discovery via a Recurrent Mapping Model</i> Yuhang Bai, Richong Zhang, Fanshuang Kong, Junfan Chen and Yongyi Mao .....	2912

<i>Modeling the Influence of Verb Aspect on the Activation of Typical Event Locations with BERT</i> Won Ik Cho, Emmanuele Chersoni, Yu-Yin Hsu and Chu-Ren Huang .....	2922
<i>On the Interaction of Belief Bias and Explanations</i> Ana Valeria González, Anna Rogers and Anders Søgaard .....	2930
<i>Combining Static Word Embeddings and Contextual Representations for Bilingual Lexicon Induction</i> Jinpeng Zhang, Baijun Ji, Nini Xiao, Xiangyu Duan, Min Zhang, Yangbin Shi and Weihua Luo	2943
<i>Exploring Unsupervised Pretraining Objectives for Machine Translation</i> Christos Baziotis, Ivan Titov, Alexandra Birch and Barry Haddow .....	2956
<i>Knowledge-Grounded Dialogue Generation with Term-level De-noising</i> Wen Zheng, Natasa Milic-Frayling and Ke Zhou .....	2972
<i>Inspecting the concept knowledge graph encoded by modern language models</i> Carlos Aspillaga, Marcelo Mendoza and Alvaro Soto .....	2984
<i>Language Tags Matter for Zero-Shot Neural Machine Translation</i> Liwei Wu, Shanbo Cheng, Mingxuan Wang and Lei Li .....	3001
<i>Latent Reasoning for Low-Resource Question Generation</i> Xinting Huang, Jianzhong Qi, Yu Sun and Rui Zhang .....	3008
<i>Probing Pre-Trained Language Models for Disease Knowledge</i> Israa Alghanmi, Luis Espinosa Anke and Steven Schockaert .....	3023
<i>AugVic: Exploiting BiText Vicinity for Low-Resource NMT</i> Tasnim Mohiuddin, M Saiful Bari and Shafiq Joty .....	3034
<i>Provably Secure Generative Linguistic Steganography</i> Siyu Zhang, Zhongliang Yang, Jinshuai Yang and Yongfeng Huang .....	3046
<i>Retrieval Enhanced Model for Commonsense Generation</i> Han Wang, Yang Liu, Chenguang Zhu, Linjun Shou, Ming Gong, Yichong Xu and Michael Zeng	3056
<i>Decoupled Dialogue Modeling and Semantic Parsing for Multi-Turn Text-to-SQL</i> Zhi Chen, Lu Chen, Hanqi Li, Ruisheng Cao, Da Ma, Mengyue Wu and Kai Yu .....	3063
<i>Adjacency List Oriented Relational Fact Extraction via Adaptive Multi-task Learning</i> Fubang Zhao, Zhuoren Jiang, Yangyang Kang, Changlong Sun and Xiaozhong Liu .....	3075
<i>Self-Supervised Document Similarity Ranking via Contextualized Language Models and Hierarchical Inference</i> Dvir Ginzburg, Itzik Malkiel, Oren Barkan, Avi Caciularu and Noam Koenigstein .....	3088
<i>How Good Is NLP? A Sober Look at NLP Tasks through the Lens of Social Impact</i> Zhijing Jin, Geeticka Chauhan, Brian Tse, Mrinmaya Sachan and Rada Mihalcea .....	3099
<i>IgSEG: Image-guided Story Ending Generation</i> Qingbao Huang, Chuan Huang, Linzhang Mo, Jielong Wei, Yi Cai, Ho-fung Leung and Qing Li	3114

<i>Improve Query Focused Abstractive Summarization by Incorporating Answer Relevance</i> Dan Su, Tiezheng Yu and Pascale Fung .....	3124
<i>Learning a Reversible Embedding Mapping using Bi-Directional Manifold Alignment</i> Ashwinkumar Ganesan, Francis Ferraro and Tim Oates .....	3132
<i>Probabilistic Graph Reasoning for Natural Proof Generation</i> Changzhi Sun, Xinbo Zhang, Jiangjie Chen, Chun Gan, Yuanbin Wu, Jiaze Chen, Hao Zhou and Lei Li .....	3140
<i>Enhancing Zero-shot and Few-shot Stance Detection with Commonsense Knowledge Graph</i> Rui Liu, Zheng Lin, Yutong Tan and Weiping Wang .....	3152
<i>Dialogue Graph Modeling for Conversational Machine Reading</i> Siru Ouyang, Zhuosheng Zhang and Hai Zhao .....	3158
<i>IndoCollex: A Testbed for Morphological Transformation of Indonesian Word Colloquialism</i> Haryo Akbarianto Wibowo, Made Nindyatama Nityasya, Afra Feyza Akyürek, Suci Fitriany, Alham Fikri Aji, Radityo Eko Prasojito and Derry Tanti Wijaya .....	3170
<i>Manifold Adversarial Augmentation for Neural Machine Translation</i> Guandan Chen, Kai Fan, Kaibo Zhang, Boxing Chen and Zhongqiang Huang .....	3184
<i>Learning to Bridge Metric Spaces: Few-shot Joint Learning of Intent Detection and Slot Filling</i> Yutai Hou, Yongkui Lai, cheng chen, Wanxiang Che and Ting Liu .....	3190
<i>Insertion-based Tree Decoding</i> Denis Lukovnikov and Asja Fischer .....	3201
<i>Is the Lottery Fair? Evaluating Winning Tickets Across Demographics</i> Victor Petrén Bach Hansen and Anders Søgaard .....	3214
<i>SSMix: Saliency-Based Span Mixup for Text Classification</i> Soyoung Yoon, Gyuwan Kim and Kyumin Park .....	3225
<i>Detecting Bot-Generated Text by Characterizing Linguistic Accommodation in Human-Bot Interactions</i> Paras Bhatt and Anthony Rios .....	3235
<i>Defending Pre-trained Language Models from Adversarial Word Substitution Without Performance Sac- rifice</i> Rongzhou Bao, Jiayi Wang and Hai Zhao .....	3248
<i>BERT-Proof Syntactic Structures: Investigating Errors in Discontinuous Constituency Parsing</i> Maximin Coavoux .....	3259
<i>DoT: An efficient Double Transformer for NLP tasks with tables</i> Syrine Krichene, Thomas Müller and Julian Eisenschlos .....	3273
<i>Grammatical Error Correction as GAN-like Sequence Labeling</i> Kevin Parnow, Zuchao Li and Hai Zhao .....	3284
<i>Neural Entity Recognition with Gazetteer based Fusion</i> Qing Sun and Parminder Bhatia .....	3291
<i>Hyperbolic Temporal Knowledge Graph Embeddings with Relational and Time Curvatures</i> Sebastien Montella, Lina M. Rojas Barahona and Johannes Heinecke .....	3296

<i>Disfl-QA: A Benchmark Dataset for Understanding Disfluencies in Question Answering</i> Aditya Gupta, Jiacheng Xu, Shyam Upadhyay, Diyi Yang and Manaal Faruqui . . . . .	3309
<i>Does Robustness Improve Fairness? Approaching Fairness with Word Substitution Robustness Methods for Text Classification</i> Yada Pruksachatkun, Satyapriya Krishna, Jwala Dhamala, Rahul Gupta and Kai-Wei Chang .	3320
<i>A Joint Model for Structure-based News Genre Classification with Application to Text Summarization</i> Zeyu Dai and Ruihong Huang . . . . .	3332
<i>Representing Syntax and Composition with Geometric Transformations</i> Lorenzo Bertolini, Julie Weeds, David Weir and Qiwei Peng . . . . .	3343
<i>Figurative Language in Recognizing Textual Entailment</i> Tuhin Chakraborty, Debanjan Ghosh, Adam Poliak and Smaranda Muresan . . . . .	3354
<i>To Point or Not to Point: Understanding How Abstractive Summarizers Paraphrase Text</i> Matt Wilber, William Timkey and Marten van Schijndel . . . . .	3362
<i>AgreeSum: Agreement-Oriented Multi-Document Summarization</i> Richard Yuanzhe Pang, Adam Lelkes, Vinh Tran and Cong Yu . . . . .	3377
<i>BERT Busters: Outlier Dimensions that Disrupt Transformers</i> Olga Kovaleva, Saurabh Kulshreshtha, Anna Rogers and Anna Rumshisky . . . . .	3392
<i>"We will Reduce Taxes" - Identifying Election Pledges with Language Models</i> Tommaso Fornaciari, Dirk Hovy, Elin Naurin, Julia Runeson, Robert Thomson and Pankaj Adhikari	3406
<i>WeaQA: Weak Supervision via Captions for Visual Question Answering</i> Pratyay Banerjee, Tejas Gokhale, Yezhou Yang and Chitta Baral . . . . .	3420
<i>How well do you know your summarization datasets?</i> Priyam Tejaswin, Dhruv Naik and Pengfei Liu . . . . .	3436
<i>Multilingual Translation from Denoising Pre-Training</i> Yuqing Tang, Chau Tran, Xian Li, Peng-Jen Chen, Naman Goyal, Vishrav Chaudhary, Jiatao Gu and Angela Fan . . . . .	3450
<i>Annotations Matter: Leveraging Multi-task Learning to Parse UD and SUD</i> Zeeshan Ali Sayyed and Daniel Dakota . . . . .	3467
<i>Generating Informative Conclusions for Argumentative Texts</i> Shahbaz Syed, Khalid Al Khatib, Milad Alshomary, Henning Wachsmuth and Martin Potthast	3482
<i>Substructure Substitution: Structured Data Augmentation for NLP</i> Haoyue Shi, Karen Livescu and Kevin Gimpel . . . . .	3494
<i>Towards Protecting Vital Healthcare Programs by Extracting Actionable Knowledge from Policy</i> Vanessa Lopez, Nagesh Yadav, Gabriele Picco, Inge Vejsbjerg, Eoin Carrol, Seamus Brady, Marco Luca Sbodio, Lam Thanh Hoang, Miao Wei and John Segrave . . . . .	3509
<i>Not Far Away, Not So Close: Sample Efficient Nearest Neighbour Data Augmentation via MiniMax</i> Ehsan Kamalloo, Mehdi Rezagholizadeh, Peyman Passban and Ali Ghodsi . . . . .	3522

<i>It's All in the Heads: Using Attention Heads as a Baseline for Cross-Lingual Transfer in Commonsense Reasoning</i>	
Alexey Tikhonov and Max Ryabinin .....	3534
<i>Biomedical Interpretable Entity Representations</i>	
Diego Garcia-Olano, Yasumasa Onoe, Ioana Baldini, Joydeep Ghosh, Byron Wallace and Kush Varshney .....	3547
<i>Learning Robust Latent Representations for Controllable Speech Synthesis</i>	
Shakti Kumar, Jithin Pradeep and Hussain Zaidi .....	3562
<i>How to Split: the Effect of Word Segmentation on Gender Bias in Speech Translation</i>	
Marco Gaido, Beatrice Savoldi, Luisa Bentivogli, Matteo Negri and Marco Turchi .....	3576
<i>On the Ethical Limits of Natural Language Processing on Legal Text</i>	
Dimitrios Tsarapatsanis and Nikolaos Aletras .....	3590
<i>An Exploratory Analysis of the Relation between Offensive Language and Mental Health</i>	
Ana-Maria Bucur, Marcos Zampieri and Liviu P. Dinu .....	3600
<i>Transforming Term Extraction: Transformer-Based Approaches to Multilingual Term Extraction Across Domains</i>	
Christian Lang, Lennart Wachowiak, Barbara Heinisch and Dagmar Gromann .....	3607
<i>ProofWriter: Generating Implications, Proofs, and Abductive Statements over Natural Language</i>	
Oyvind Tafjord, Bhavana Dalvi and Peter Clark .....	3621
<i>Probing Image-Language Transformers for Verb Understanding</i>	
Lisa Anne Hendricks and Aida Nematzadeh .....	3635
<i>Implications of Using Internet Sting Corpora to Approximate Underage Victims</i>	
Tatiana Ringenber, Kathryn Seigfried-Spellar and Julia Rayz .....	3645
<i>Detecting Domain Polarity-Changes of Words in a Sentiment Lexicon</i>	
Shuai Wang, Guangyi Lv, Sahisnu Mazumder and Bing Liu .....	3657
<i>Analyzing Online Political Advertisements</i>	
Danae Sánchez Villegas, Saeid Mokaram and Nikolaos Aletras .....	3669
<i>Do Language Models Perform Generalizable Commonsense Inference?</i>	
Peifeng Wang, Filip Ilievski, Muhao Chen and Xiang Ren .....	3681
<i>Probing Multi-modal Machine Translation with Pre-trained Language Model</i>	
Kong Yawei and Kai Fan .....	3689
<i>The interplay between language similarity and script on a novel multi-layer Algerian dialect corpus</i>	
Samia Touileb and Jeremy Barnes .....	3700
<i>Few-Shot Upsampling for Protest Size Detection</i>	
Andrew Halterman and Benjamin J. Radford .....	3713
<i>Modeling the Unigram Distribution</i>	
Irene Nikkarinen, Tiago Pimentel, Damián Blasi and Ryan Cotterell .....	3721

<i>On the Lack of Robust Interpretability of Neural Text Classifiers</i>	
Muhammad Bilal Zafar, Michele Donini, Dylan Slack, Cedric Archambeau, Sanjiv Das and Krishnaram Kenthapadi . . . . .	3730
<i>Multimodal Graph-based Transformer Framework for Biomedical Relation Extraction</i>	
Sriram Pingali, Shweta Yadav, Pratik Dutta and Sriparna Saha . . . . .	3741
<i>Summary Grounded Conversation Generation</i>	
Chulaka Gunasekara, Guy Feigenblat, Benjamin Sznajder, Sachindra Joshi and David Konopnicki	3748
<i>A Non-Autoregressive Edit-Based Approach to Controllable Text Simplification</i>	
Sweta Agrawal, Weijia Xu and Marine Carpuat . . . . .	3757
<i>Investigating Transfer Learning in Multilingual Pre-trained Language Models through Chinese Natural Language Inference</i>	
Hai Hu, He Zhou, Zuoyu Tian, Yiwen Zhang, Yina Patterson, Yanting Li, Yixin Nie and Kyle Richardson . . . . .	3770
<i>Using surprisal and fMRI to map the neural bases of broad and local contextual prediction during natural language comprehension</i>	
Shohini Bhattachali and Philip Resnik . . . . .	3786
<i>Assessing the Syntactic Capabilities of Transformer-based Multilingual Language Models</i>	
Laura Pérez-Mayos, Alba Táboas García, Simon Mille and Leo Wanner . . . . .	3799
<i>Are Larger Pretrained Language Models Uniformly Better? Comparing Performance at the Instance Level</i>	
Ruiqi Zhong, Dhruva Ghosh, Dan Klein and Jacob Steinhardt . . . . .	3813
<i>Named Entity Recognition through Deep Representation Learning and Weak Supervision</i>	
Jerrod Parker and Shi Yu . . . . .	3828
<i>Explaining NLP Models via Minimal Contrastive Editing (MiCE)</i>	
Alexis Ross, Ana Marasović and Matthew Peters . . . . .	3840
<i>Differential Privacy for Text Analytics via Natural Text Sanitization</i>	
Xiang Yue, Minxin Du, Tianhao Wang, Yaliang Li, Huan Sun and Sherman S. M. Chow . . . . .	3853
<i>Synthesizing Adversarial Negative Responses for Robust Response Ranking and Evaluation</i>	
Prakhar Gupta, Yulia Tsvetkov and Jeffrey Bigham . . . . .	3867
<i>Leveraging Abstract Meaning Representation for Knowledge Base Question Answering</i>	
Pavan Kapanipathi, Ibrahim Abdelaziz, Srinivas Ravishankar, Salim Roukos, Alexander Gray, Ramón Fernandez Astudillo, Maria Chang, Cristina Cornelio, Saswati Dana, Achille Fokoue, Dinesh Garg, Alfio Gliozzo, Sairam Gurajada, Hima Karanam, Naweed Khan, Dinesh Khandelwal, Young-Suk Lee, Yunyao Li, Francois Luus, Ndivhuwo Makondo, Nandana Mihindukulasooriya, Tahira Naseem, Sumit Neelam, Lucian Popa, Revanth Gangi Reddy, Ryan Riegel, Gaetano Rossiello, Udit Sharma, G P Shrivatsa Bhargav and Mo Yu . . . . .	3884
<i>On the Gap between Adoption and Understanding in NLP</i>	
Federico Bianchi and Dirk Hovy . . . . .	3895
<i>Learning Disentangled Latent Topics for Twitter Rumour Veracity Classification</i>	
John Dougrez-Lewis, Maria Liakata, Elena Kochkina and Yulan He . . . . .	3902

<i>Perceptual Models of Machine-Edited Text</i>	
Elizabeth Merkhofer, Monica-Ann Mendoza, Rebecca Marvin and John Henderson . . . . .	3909
<i>Scaling Within Document Coreference to Long Texts</i>	
Raghuv eer Thirukovalluru, Nicholas Monath, Kumar Shridhar, Manzil Zaheer, Mrinmaya Sachan and Andrew McCallum . . . . .	3921
<i>LEWIS: Levenshtein Editing for Unsupervised Text Style Transfer</i>	
Machel Reid and Victor Zhong . . . . .	3932
<i>Constructing Flow Graphs from Procedural Cybersecurity Texts</i>	
Kuntal Kumar Pal, Kazuaki Kashihara, Pratyay Banerjee, Swaroop Mishra, Ruoyu Wang and Chitta Baral . . . . .	3945
<i>Cluster-Former: Clustering-based Sparse Transformer for Question Answering</i>	
Shuohang Wang, Luowei Zhou, Zhe Gan, Yen-Chun Chen, Yuwei Fang, Siqi Sun, Yu Cheng and Jingjing Liu . . . . .	3958
<i>Minimally-Supervised Morphological Segmentation using Adaptor Grammars with Linguistic Priors</i>	
Ramy Eskander, Cass Lowry, Sujay Khandagale, Francesca Callejas, Judith Klavans, Maria Polinsky and Smaranda Muresan . . . . .	3969
<i>Multi-Task Learning and Adapted Knowledge Models for Emotion-Cause Extraction</i>	
Elsbeth Turcan, Shuai Wang, Rishita Anubhai, Kasturi Bhattacharjee, Yaser Al-Onaizan and Smaranda Muresan . . . . .	3975
<i>The Utility and Interplay of Gazetteers and Entity Segmentation for Named Entity Recognition in English</i>	
Oshin Agarwal and Ani Nenkova . . . . .	3990
<i>On the Cost-Effectiveness of Stacking of Neural and Non-Neural Methods for Text Classification: Scenarios and Performance Prediction</i>	
Christian Gomes, Marcos Goncalves, Leonardo Rocha and Sergio Canuto . . . . .	4003
<i>Unsupervised Domain Adaptation for Event Detection using Domain-specific Adapters</i>	
Nghia Ngo Trung, Duy Phung and Thien Huu Nguyen . . . . .	4015
<i>Predicting in-hospital mortality by combining clinical notes with time-series data</i>	
Iman Deznabi, Mohit Iyyer and Madalina Fiterau . . . . .	4026
<i>Sequence Models for Computational Etymology of Borrowings</i>	
Winston Wu, Kevin Duh and David Yarowsky . . . . .	4032
<i>Learning Contextualized Knowledge Structures for Commonsense Reasoning</i>	
Jun Yan, Mrigank Raman, Aaron Chan, Tianyu Zhang, Ryan Rossi, Handong Zhao, Sungchul Kim, Nedim Lipka and Xiang Ren . . . . .	4038
<i>Analyzing Stereotypes in Generative Text Inference Tasks</i>	
Anna Sotnikova, Yang Trista Cao, Hal Daumé III and Rachel Rudinger . . . . .	4052
<i>HySPA: Hybrid Span Generation for Scalable Text-to-Graph Extraction</i>	
Liliang Ren, Chenkai Sun, Heng Ji and Julia Hockenmaier . . . . .	4066
<i>Improving Automated Evaluation of Open Domain Dialog via Diverse Reference Augmentation</i>	
Varun Gangal, Harsh Jhamtani, Eduard Hovy and Taylor Berg-Kirkpatrick . . . . .	4079

<i>Who Blames or Endorses Whom? Entity-to-Entity Directed Sentiment Extraction in News Text</i> Kunwoo Park, Zhufeng Pan and Jungseock Joo .....	4091
<i>New Dataset and Strong Baselines for the Grammatical Error Correction of Russian</i> Viet Anh Trinh and Alla Rozovskaya .....	4103
<i>A Formidable Ability: Detecting Adjectival Extremeness with DSMs</i> Farhan Samir, Barend Beekhuizen and Suzanne Stevenson.....	4112
<i>Effective Attention Sheds Light On Interpretability</i> Kaiser Sun and Ana Marasović .....	4126
<i>Compositionality of Complex Graphemes in the Undeciphered Proto-Elamite Script using Image and Text Embedding Models</i> Logan Born, Kathryn Kelley, M. Willis Monroe and Anoop Sarkar .....	4136
<i>On the Distribution, Sparsity, and Inference-time Quantization of Attention Values in Transformers</i> Tianchu Ji, Shraddhan Jain, Michael Ferdman, Peter Milder, H. Andrew Schwartz and Niranjan Balasubramanian.....	4147
<i>Ethical-Advice Taker: Do Language Models Understand Natural Language Interventions?</i> Jieyu Zhao, Daniel Khashabi, Tushar Khot, Ashish Sabharwal and Kai-Wei Chang .....	4158
<i>Unsupervised Label Refinement Improves Dataless Text Classification</i> Zewei Chu, Karl Stratos and Kevin Gimpel .....	4165
<i>Prompting Contrastive Explanations for Commonsense Reasoning Tasks</i> Bhargavi Paranjape, Julian Michael, Marjan Ghazvininejad, Hannaneh Hajishirzi and Luke Zettlemoyer.....	4179
<i>SMS Spam Detection Through Skip-gram Embeddings and Shallow Networks</i> Gustavo Sousa, Daniel Carlos Guimarães Pedronette, João Paulo Papa and Ivan Rizzo Guilherme	4193
<i>Hierarchical Task Learning from Language Instructions with Unified Transformers and Self-Monitoring</i> Yichi Zhang and Joyce Chai .....	4202
<i>Marked Attribute Bias in Natural Language Inference</i> Hillary Dawkins.....	4214
<i>VLM: Task-agnostic Video-Language Model Pre-training for Video Understanding</i> Hu Xu, Gargi Ghosh, Po-Yao Huang, Prahal Arora, Masoumeh Aminzadeh, Christoph Feichtenhofer, Florian Metze and Luke Zettlemoyer .....	4227
<i>Corpus-Level Evaluation for Event QA: The IndiaPoliceEvents Corpus Covering the 2002 Gujarat Violence</i> Andrew Halterman, Katherine Keith, Sheikh Sarwar and Brendan O'Connor .....	4240
<i>Memory-Efficient Differentiable Transformer Architecture Search</i> Yuekai Zhao, Li Dong, Yelong Shen, Zhihua Zhang, Furu Wei and Weizhu Chen .....	4254
<i>On the Copying Behaviors of Pre-Training for Neural Machine Translation</i> Xuebo Liu, Longyue Wang, Derek F. Wong, Liang Ding, Lidia S. Chao, Shuming Shi and Zhaopeng Tu .....	4265



<i>Answer Generation for Retrieval-based Question Answering Systems</i>	
Chao-Chun Hsu, Eric Lind, Luca Soldaini and Alessandro Moschitti .....	4276
<i>Grounding 'Grounding' in NLP</i>	
Khyathi Raghavi Chandu, Yonatan Bisk and Alan W Black .....	4283
<i>Federated Chinese Word Segmentation with Global Character Associations</i>	
Yuanhe Tian, Guimin Chen, Han Qin and Yan Song .....	4306
<i>PSED: A Dataset for Selecting Emphasis in Presentation Slides</i>	
Amirreza Shirani, Giai Tran, Hieu Trinh, Franck Dernoncourt, Nedim Lipka, Jose Echevarria, Thamar Solorio and Paul Asente .....	4314
<i>MLMLM: Link Prediction with Mean Likelihood Masked Language Model</i>	
Louis Clouatre, Philippe Trempe, Amal Zouaq and Sarath Chandar .....	4321
<i>Modulating Language Models with Emotions</i>	
Ruibao Liu, Jason Wei, Chenyan Jia and Soroush Vosoughi .....	4332
<i>Effective Batching for Recurrent Neural Network Grammars</i>	
Hiroshi Noji and Yohei Oseki .....	4340
<i>Verb Sense Clustering using Contextualized Word Representations for Semantic Frame Induction</i>	
Kosuke Yamada, Ryohei Sasano and Koichi Takeda .....	4353
<i>Benchmarking Neural Topic Models: An Empirical Study</i>	
Thanh-Nam Doan and Tuan-Anh Hoang .....	4363
<i>Enhancing Chinese Word Segmentation via Pseudo Labels for Practicability</i>	
Kaiyu Huang, Junpeng Liu, Degen Huang, Deyi Xiong, Zhuang Liu and Jinsong Su .....	4369
<i>Analysis of Tree-Structured Architectures for Code Generation</i>	
Samip Dahal, Adyasha Maharana and Mohit Bansal .....	4382
<i>How Does Distilled Data Complexity Impact the Quality and Confidence of Non-Autoregressive Machine Translation?</i>	
Weijia Xu, Shuming Ma, Dongdong Zhang and Marine Carpuat .....	4392
<i>Leveraging Topic Relatedness for Argument Persuasion</i>	
Xinran Zhao, Esin Durmus, Hongming Zhang and Claire Cardie .....	4401
<i>One Teacher is Enough? Pre-trained Language Model Distillation from Multiple Teachers</i>	
Chuhan Wu, Fangzhao Wu and Yongfeng Huang .....	4408
<i>Logic-Consistency Text Generation from Semantic Parses</i>	
Chang Shu, Yusen Zhang, Xiangyu Dong, Peng Shi, Tao Yu and Rui Zhang .....	4414
<i>Inducing Semantic Roles Without Syntax</i>	
Julian Michael and Luke Zettlemoyer .....	4427
<i>Plot and Rework: Modeling Storylines for Visual Storytelling</i>	
Chi-yang Hsu, Yun-Wei Chu, Ting-Hao Huang and Lun-Wei Ku .....	4443
<i>Disentangled Code Representation Learning for Multiple Programming Languages</i>	
Jingfeng Zhang, Haiwen Hong, Yin Zhang, Yao Wan, Ye Liu and Yulei Sui .....	4454

<i>Exploring Self-Identified Counseling Expertise in Online Support Forums</i>	
Allison Lahnala, Yuntian Zhao, Charles Welch, Jonathan K. Kummerfeld, Lawrence C An, Kenneth Resnicow, Rada Mihalcea and Verónica Pérez-Rosas . . . . .	4467
<i>An Investigation of Suitability of Pre-Trained Language Models for Dialogue Generation – Avoiding Discrepancies</i>	
Yan Zeng and Jian-Yun Nie . . . . .	4481
<i>Learning to Sample Replacements for ELECTRA Pre-Training</i>	
Yaru Hao, Li Dong, Hangbo Bao, Ke Xu and Furu Wei . . . . .	4495
<i>Reordering Examples Helps during Priming-based Few-Shot Learning</i>	
Sawan Kumar and Partha Talukdar . . . . .	4507
<i>Constrained Labeled Data Generation for Low-Resource Named Entity Recognition</i>	
Ruohao Guo and Dan Roth . . . . .	4519
<i>He is very intelligent, she is very beautiful? On Mitigating Social Biases in Language Modelling and Generation</i>	
Aparna Garimella, Akhash Amarnath, Kiran Kumar, Akash Pramod Yalla, Anandhavelu N, Niyati Chhaya and Balaji Vasani Srinivasan . . . . .	4534
<i>Task-adaptive Pre-training of Language Models with Word Embedding Regularization</i>	
Kosuke Nishida, Kyosuke Nishida and Sen Yoshida . . . . .	4546
<i>Do Grammatical Error Correction Models Realize Grammatical Generalization?</i>	
Masato Mita and Hitomi Yanaka . . . . .	4554
<i>Domain-Aware Dependency Parsing for Questions</i>	
Aparna Garimella, Laura Chiticariu and Yunyao Li . . . . .	4562
<i>Using Social and Linguistic Information to Adapt Pretrained Representations for Political Perspective Identification</i>	
Chang Li and Dan Goldwasser . . . . .	4569
<i>Enhancing Dialogue-based Relation Extraction by Speaker and Trigger Words Prediction</i>	
Tianyang Zhao, Zhao Yan, Yunbo Cao and Zhoujun Li . . . . .	4580
<i>Modeling Event-Pair Relations in External Knowledge Graphs for Script Reasoning</i>	
Yucheng Zhou, Xiubo Geng, Tao Shen, Jian Pei, Wenqiang Zhang and Daxin Jiang . . . . .	4586
<i>PROST: Physical Reasoning about Objects through Space and Time</i>	
Stéphane Aroca-Ouellette, Cory Paik, Alessandro Roncone and Katharina Kann . . . . .	4597
<i>Revisiting the Evaluation of End-to-end Event Extraction</i>	
Shun Zheng, Wei Cao, Wei Xu and Jiang Bian . . . . .	4609
<i>Direct Simultaneous Speech-to-Text Translation Assisted by Synchronized Streaming ASR</i>	
Junkun Chen, Mingbo Ma, Renjie Zheng and Liang Huang . . . . .	4618
<i>HIT - A Hierarchically Fused Deep Attention Network for Robust Code-mixed Language Representation</i>	
Ayan Sengupta, Sourabh Kumar Bhattacharjee, Tanmoy Chakraborty and Md. Shad Akhtar . . . . .	4625
<i>Semi-Supervised Data Programming with Subset Selection</i>	
Ayush Maheshwari, Oishik Chatterjee, Krishnateja Killamsetty, Ganesh Ramakrishnan and Rishabh Iyer . . . . .	4640

<i>Fingerprinting Fine-tuned Language Models in the Wild</i>	
Nirav Diwan, Tanmoy Chakraborty and Zubair Shafiq .....	4652
<i>Analyzing Code Embeddings for Coding Clinical Narratives</i>	
Wei Shi, Jiewen Wu, Xiwen Yang, Nancy Chen, Ivan Ho Mien, Jung-Jae Kim and Pavitra Krishnaswamy .....	4665
<i>Automatic Construction of Sememe Knowledge Bases via Dictionaries</i>	
Fanchao Qi, Yangyi Chen, Fengyu Wang, Zhiyuan Liu, Xiao Chen and Maosong Sun .....	4673
<i>Rule-Aware Reinforcement Learning for Knowledge Graph Reasoning</i>	
Zhongni Hou, Xiaolong Jin, Zixuan Li and Long Bai .....	4687
<i>XL-Sum: Large-Scale Multilingual Abstractive Summarization for 44 Languages</i>	
Tahmid Hasan, Abhik Bhattacharjee, Md. Saiful Islam, Kazi Mubasshir, Yuan-Fang Li, Yong-Bin Kang, M. Sohel Rahman and Rifat Shahriyar .....	4693
<i>Use of Formal Ethical Reviews in NLP Literature: Historical Trends and Current Practices</i>	
Sebastin Santy, Anku Rani and Monojit Choudhury .....	4704
<i>As Easy as 1, 2, 3: Behavioural Testing of NMT Systems for Numerical Translation</i>	
Jun Wang, Chang Xu, Francisco Guzmán, Ahmed El-Kishky, Benjamin Rubinstein and Trevor Cohn .....	4711
<i>Investigating Memorization of Conspiracy Theories in Text Generation</i>	
Sharon Levy, Michael Saxon and William Yang Wang .....	4718
<i>A Text-Centered Shared-Private Framework via Cross-Modal Prediction for Multimodal Sentiment Analysis</i>	
Yang Wu, Zijie Lin, Yanyan Zhao, Bing Qin and Li-Nan ZHU .....	4730
<i>What Would a Teacher Do? Predicting Future Talk Moves</i>	
Ananya Ganesh, Martha Palmer and Katharina Kann .....	4739
<i>BioGen: Generating Biography Summary under Table Guidance on Wikipedia</i>	
Shen Gao, Xiuying Chen, Chang Liu, Dongyan Zhao and Rui Yan .....	4752
<i>Multilingual Simultaneous Neural Machine Translation</i>	
Philip Arthur, Dongwon Ryu and Gholamreza Haffari .....	4758
<i>Cross-Domain Review Generation for Aspect-Based Sentiment Analysis</i>	
Jianfei Yu, Chenggong Gong and Rui Xia .....	4767
<i>On the Language Coverage Bias for Neural Machine Translation</i>	
Shuo Wang, Zhaopeng Tu, Zhixing Tan, Shuming Shi, Maosong Sun and Yang Liu .....	4778
<i>Named Entity Recognition via Noise Aware Training Mechanism with Data Filter</i>	
Xiusheng Huang, Yubo Chen, Shun Wu, Jun Zhao, Yuantao Xie and Weijian Sun .....	4791
<i>A Multi-Task Approach for Improving Biomedical Named Entity Recognition by Incorporating Multi-Granularity information</i>	
Yiqi Tong, Yidong Chen and xiaodong shi .....	4804
<i>EBERT: Efficient BERT Inference with Dynamic Structured Pruning</i>	
Zejian Liu, Fanrong Li, Gang Li and Jian Cheng .....	4814

<i>Strong and Light Baseline Models for Fact-Checking Joint Inference</i> Kateryna Tymoshenko and Alessandro Moschitti .....	4824
<i>Sketch and Refine: Towards Faithful and Informative Table-to-Text Generation</i> Peng Wang, Junyang Lin, An Yang, Chang Zhou, Yichang Zhang, Jingren Zhou and Hongxia Yang	4831
<i>TILGAN: Transformer-based Implicit Latent GAN for Diverse and Coherent Text Generation</i> Shizhe Diao, Xinwei Shen, Kashun Shum, Yan Song and Tong Zhang .....	4844
<i>John praised Mary because _he_? Implicit Causality Bias and Its Interaction with Explicit Cues in LMs</i> Yova Kementchedjheva, Mark Anderson and Anders Søgaard .....	4859
<i>Do It Once: An Embarrassingly Simple Joint Matching Approach to Response Selection</i> Linhao Zhang, Dehong Ma, Sujian Li and Houfeng Wang .....	4872
<i>Climbing the Tower of Treebanks: Improving Low-Resource Dependency Parsing via Hierarchical Source Selection</i> Goran Glavaš and Ivan Vulić .....	4878
<i>Enhancing the Open-Domain Dialogue Evaluation in Latent Space</i> Zhangming Chan, Lemao Liu, Juntao Li, Haisong Zhang, Dongyan Zhao, Shuming Shi and Rui Yan .....	4889
<i>Adapting Monolingual Models: Data can be Scarce when Language Similarity is High</i> Wietse de Vries, Martijn Bartelds, Malvina Nissim and Martijn Wieling .....	4901
<i>BatchMixup: Improving Training by Interpolating Hidden States of the Entire Mini-batch</i> Wenpeng Yin, Huan Wang, Jin Qu and Caiming Xiong .....	4908
<i>DocNLI: A Large-scale Dataset for Document-level Natural Language Inference</i> Wenpeng Yin, Dragomir Radev and Caiming Xiong .....	4913
<i>Rule Augmented Unsupervised Constituency Parsing</i> Atul Sahay, Anshul Nasery, Ayush Maheshwari, Ganesh Ramakrishnan and Rishabh Iyer ....	4923
<i>Are Multilingual Models the Best Choice for Moderately Under-resourced Languages? A Comprehensive Assessment for Catalan</i> Jordi Armengol-Estapé, Casimiro Pio Carrino, Carlos Rodriguez-Penagos, Ona de Gibert Bonet, Carme Armentano-Oller, Aitor Gonzalez-Agirre, Maite Melero and Marta Villegas .....	4933
<i>How transfer learning impacts linguistic knowledge in deep NLP models?</i> Nadir Durrani, Hassan Sajjad and Fahim Dalvi .....	4947
<i>Language Models Use Monotonicity to Assess NPI Licensing</i> Jaap Jumelet, Milica Denic, Jakub Szymanik, Dieuwke Hupkes and Shane Steinert-Threlkeld	4958
<i>Slot Transferability for Cross-domain Slot Filling</i> Hengtong Lu, zhuoxin han, Caixia Yuan, Xiaojie WANG, Shuyu Lei, Huixing Jiang and Wei Wu	4970
<i>Word Graph Guided Summarization for Radiology Findings</i> jinpeng hu, Jianling Li, Zhihong Chen, Yaling Shen, Yan Song, Xiang Wan and Tsung-Hui Chang	4980

<i>Generalized Supervised Attention for Text Generation</i>	
Yixian Liu, Liwen Zhang, Xinyu Zhang, Yong Jiang, Yue Zhang and Kewei Tu . . . . .	4991
<i>Uncertainty Aware Review Hallucination for Science Article Classification</i>	
Korbinian Friedl, Georgios Rizos, Lukas Stappen, Madina Hasan, Lucia Specia, Thomas Hain and Björn Schuller . . . . .	5004
<i>Automatically Select Emotion for Response via Personality-affected Emotion Transition</i>	
Zhiyuan Wen, Jiannong Cao, Ruosong Yang, Shuaiqi LIU and Jiaying Shen . . . . .	5010
<i>Highlight-Transformer: Leveraging Key Phrase Aware Attention to Improve Abstractive Multi-Document Summarization</i>	
Shuaiqi LIU, Jiannong Cao, Ruosong Yang and Zhiyuan Wen . . . . .	5021
<i>Phrase-Level Action Reinforcement Learning for Neural Dialog Response Generation</i>	
Takato Yamazaki and Akiko Aizawa . . . . .	5028
<i>Automatic Speech Recognition in Sanskrit: A New Speech Corpus and Modelling Insights</i>	
Devaraja Adiga, Rishabh Kumar, Amrith Krishna, Preethi Jyothi, Ganesh Ramakrishnan and Pawan Goyal . . . . .	5039
<i>Constraint based Knowledge Base Distillation in End-to-End Task Oriented Dialogs</i>	
Dinesh Raghu, Atishya Jain, and Sachindra Joshi . . . . .	5051
<i>DialogSum: A Real-Life Scenario Dialogue Summarization Dataset</i>	
Yulong Chen, Yang Liu, Liang Chen and Yue Zhang . . . . .	5062
<i>What Did You Refer to? Evaluating Co-References in Dialogue</i>	
Wei-Nan Zhang, Yue Zhang, Hanlin Tang, Zhengyu Zhao, Caihai Zhu and Ting Liu . . . . .	5075
<i>Beyond Metadata: What Paper Authors Say About Corpora They Use</i>	
Nikolay Kolyada, Martin Potthast and Benno Stein . . . . .	5085
<i>Knowledge Distillation for Quality Estimation</i>	
Amit Gajbhiye, Marina Fomicheva, Fernando Alva-Manchego, Frédéric Blain, Abiola Obamuyide, Nikolaos Aletras and Lucia Specia . . . . .	5091
<i>Cross-document Coreference Resolution over Predicted Mentions</i>	
Arie Cattan, Alon Eirew, Gabriel Stanovsky, Mandar Joshi and Ido Dagan . . . . .	5100
<i>Controllable Abstractive Dialogue Summarization with Sketch Supervision</i>	
Chien-Sheng Wu, Linqing Liu, Wenhao Liu, Pontus Stenetorp and Caiming Xiong . . . . .	5108
<i>Elaborative Simplification: Content Addition and Explanation Generation in Text Simplification</i>	
Neha Srikanth and Junyi Jessy Li . . . . .	5123
<i>Could you give me a hint ? Generating inference graphs for defeasible reasoning</i>	
Aman Madaan, Dheeraj Rajagopal, Niket Tandon, Yiming Yang and Eduard Hovy . . . . .	5138
<i>Characterizing Social Spambots by their Human Traits</i>	
Salvatore Giorgi, Lyle Ungar and H. Andrew Schwartz . . . . .	5148