

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

Volume 1 (Long Papers)

Online  
1-6 August 2021

Part 1 of 11

ISBN: 978-1-7138-3308-6

**Printed from e-media with permission by:**

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



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

Copyright© (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>Investigating label suggestions for opinion mining in German Covid-19 social media</i> Tilman Beck, Ji-Ung Lee, Christina Viehmann, Marcus Maurer, Oliver Quiring and Iryna Gurevych	1
<i>How Did This Get Funded?! Automatically Identifying Quirky Scientific Achievements</i> Chen Shani, Nadav Borenstein and Dafna Shahaf	14
<i>Engage the Public: Poll Question Generation for Social Media Posts</i> Zexin Lu, Keyang Ding, Yuji Zhang, Jing Li, Baolin Peng and Lemao Liu	29
<i>HateCheck: Functional Tests for Hate Speech Detection Models</i> Paul Röttger, Bertie Vidgen, Dong Nguyen, Zeerak Waseem, Helen Margetts and Janet Pierrehumbert	41
<i>Unified Dual-view Cognitive Model for Interpretable Claim Verification</i> Lianwei Wu, Yuan Rao, Yuqian Lan, Ling Sun and Zhaoyin Qi	59
<i>DeepRapper: Neural Rap Generation with Rhyme and Rhythm Modeling</i> Lanqing Xue, Kaitao Song, Duocai Wu, Xu Tan, Nevin L. Zhang, Tao Qin, Wei-Qiang Zhang and Tie-Yan Liu	69
<i>PENS: A Dataset and Generic Framework for Personalized News Headline Generation</i> Xiang Ao, Xiting Wang, Ling Luo, Ying Qiao, Qing He and Xing Xie	82
<i>Enhancing Content Preservation in Text Style Transfer Using Reverse Attention and Conditional Layer Normalization</i> Dongkyu Lee, Zhiliang Tian, Lanqing Xue and Nevin L. Zhang	93
<i>Mention Flags (MF): Constraining Transformer-based Text Generators</i> Yufei Wang, Ian Wood, Stephen Wan, Mark Dras and Mark Johnson	103
<i>Generalising Multilingual Concept-to-Text NLG with Language Agnostic Delexicalisation</i> Giulio Zhou and Gerasimos Lampouras	114
<i>Conversations Are Not Flat: Modeling the Dynamic Information Flow across Dialogue Utterances</i> Zekang Li, Jinchao Zhang, Zhengcong Fei, Yang Feng and Jie Zhou	128
<i>Dual Slot Selector via Local Reliability Verification for Dialogue State Tracking</i> Jinyu Guo, Kai Shuang, Jijie Li and Zihan Wang	139
<i>Transferable Dialogue Systems and User Simulators</i> Bo-Hsiang Tseng, Yinpei Dai, Florian Kreyszig and Bill Byrne	152
<i>BoB: BERT Over BERT for Training Persona-based Dialogue Models from Limited Personalized Data</i> Haoyu Song, Yan Wang, Kaiyan Zhang, Wei-Nan Zhang and Ting Liu	167
<i>GL-GIN: Fast and Accurate Non-Autoregressive Model for Joint Multiple Intent Detection and Slot Filling</i> Libo Qin, Fuxuan Wei, Tianbao Xie, Xiao Xu, Wanxiang Che and Ting Liu	178
<i>Accelerating BERT Inference for Sequence Labeling via Early-Exit</i> Xiaonan Li, Yunfan Shao, Tianxiang Sun, Hang Yan, Xipeng Qiu and Xuanjing Huang	189

<i>Modularized Interaction Network for Named Entity Recognition</i>	
Fei Li, Zheng Wang, Siu Cheung Hui, Lejian Liao, Dandan Song, Jing Xu, Guoxiu He and meihuizi jia .....	200
<i>Capturing Event Argument Interaction via A Bi-Directional Entity-Level Recurrent Decoder</i>	
Xi Xiangyu, Wei Ye, Shikun Zhang, Quanxiu Wang, Huixing Jiang and Wei Wu .....	210
<i>UniRE: A Unified Label Space for Entity Relation Extraction</i>	
Yijun Wang, Changzhi Sun, Yuanbin Wu, Hao Zhou, Lei Li and Junchi Yan .....	220
<i>Refining Sample Embeddings with Relation Prototypes to Enhance Continual Relation Extraction</i>	
Li Cui, Deqing Yang, Jiabin Yu, Chengwei Hu, Jiayang Cheng, Jingjie Yi and Yanghua Xiao ..	232
<i>Contrastive Learning for Many-to-many Multilingual Neural Machine Translation</i>	
Xiao Pan, Mingxuan Wang, Liwei Wu and Lei Li .....	244
<i>Understanding the Properties of Minimum Bayes Risk Decoding in Neural Machine Translation</i>	
Mathias Müller and Rico Sennrich .....	259
<i>Multi-Head Highly Parallelized LSTM Decoder for Neural Machine Translation</i>	
Hongfei Xu, Qihui Liu, Josef van Genabith, Deyi Xiong and Meng Zhang .....	273
<i>A Bidirectional Transformer Based Alignment Model for Unsupervised Word Alignment</i>	
Jingyi Zhang and Josef van Genabith .....	283
<i>Learning Language Specific Sub-network for Multilingual Machine Translation</i>	
Zehui Lin, Liwei Wu, Mingxuan Wang and Lei Li .....	293
<i>Exploring the Efficacy of Automatically Generated Counterfactuals for Sentiment Analysis</i>	
Linyi Yang, Jiazheng Li, Pdraig Cunningham, Yue Zhang, Barry Smyth and Ruihai Dong . . .	306
<i>Bridge-Based Active Domain Adaptation for Aspect Term Extraction</i>	
Zhuang Chen and Tiejun Qian .....	317
<i>Multimodal Sentiment Detection Based on Multi-channel Graph Neural Networks</i>	
Xiaocui Yang, Shi Feng, Yifei Zhang and Daling Wang .....	328
<i>Aspect-Category-Opinion-Sentiment Quadruple Extraction with Implicit Aspects and Opinions</i>	
Hongjie Cai, Rui Xia and Jianfei Yu .....	340
<i>PASS: Perturb-and-Select Summarizer for Product Reviews</i>	
Nadav Oved and Ran Levy .....	351
<i>Deep Differential Amplifier for Extractive Summarization</i>	
Ruipeng Jia, Yanan Cao, Fang Fang, Yuchen Zhou, Zheng Fang, Yanbing Liu and Shi Wang ..	366
<i>Multi-TimeLine Summarization (MTLS): Improving Timeline Summarization by Generating Multiple Summaries</i>	
Yi Yu, Adam Jatowt, Antoine Doucet, Kazunari Sugiyama and Masatoshi Yoshikawa .....	377
<i>Self-Supervised Multimodal Opinion Summarization</i>	
Jinbae Im, Moonki Kim, Hoyeop Lee, Hyunsouk Cho and Sehee Chung .....	388
<i>A Training-free and Reference-free Summarization Evaluation Metric via Centrality-weighted Relevance and Self-referenced Redundancy</i>	
Wang Chen, Piji Li and Irwin King .....	404

<i>DESCGEN: A Distantly Supervised Dataset for Generating Entity Descriptions</i> Weijia Shi, Mandar Joshi and Luke Zettlemoyer .....	415
<i>Introducing Orthogonal Constraint in Structural Probes</i> Tomasz Limisiewicz and David Mareček .....	428
<i>Hidden Killer: Invisible Textual Backdoor Attacks with Syntactic Trigger</i> Fanchao Qi, Mukai Li, Yangyi Chen, Zhengyan Zhang, Zhiyuan Liu, Yasheng Wang and Maosong Sun .....	443
<i>Examining the Inductive Bias of Neural Language Models with Artificial Languages</i> Jennifer C. White and Ryan Cotterell .....	454
<i>Explaining Contextualization in Language Models using Visual Analytics</i> Rita Sevastjanova, Aikaterini-Lida Kalouli, Christin Beck, Hanna Schäfer and Mennatallah El-Assady .....	464
<i>Improving the Faithfulness of Attention-based Explanations with Task-specific Information for Text Classification</i> George Chrysostomou and Nikolaos Aletras .....	477
<i>Generating Landmark Navigation Instructions from Maps as a Graph-to-Text Problem</i> Raphael Schumann and Stefan Riezler .....	489
<i>E2E-VLP: End-to-End Vision-Language Pre-training Enhanced by Visual Learning</i> Haiyang Xu, Ming Yan, Chenliang Li, Bin Bi, Songfang Huang, Wenming Xiao and Fei Huang	503
<i>Learning Relation Alignment for Calibrated Cross-modal Retrieval</i> Shuhuai Ren, Junyang Lin, Guangxiang Zhao, Rui Men, An Yang, Jingren Zhou, Xu Sun and Hongxia Yang .....	514
<i>KM-BART: Knowledge Enhanced Multimodal BART for Visual Commonsense Generation</i> Yiran Xing, Zai Shi, Zhao Meng, Gerhard Lakemeyer, Yunpu Ma and Roger Wattenhofer .....	525
<i>Cascaded Head-colliding Attention</i> Lin Zheng, Zhiyong Wu and Lingpeng Kong .....	536
<i>Structural Knowledge Distillation: Tractably Distilling Information for Structured Predictor</i> Xinyu Wang, Yong Jiang, Zhaohui Yan, Zixia Jia, Nguyen Bach, Tao Wang, Zhongqiang Huang, Fei Huang and Kewei Tu .....	550
<i>Parameter-efficient Multi-task Fine-tuning for Transformers via Shared Hypernetworks</i> Rabeeh Karimi Mahabadi, Sebastian Ruder, Mostafa Dehghani and James Henderson .....	565
<i>COSY: COunterfactual SYntax for Cross-Lingual Understanding</i> SICHENG YU, Hao Zhang, Yulei Niu, Qianru Sun and Jing Jiang .....	577
<i>OoMMix: Out-of-manifold Regularization in Contextual Embedding Space for Text Classification</i> Seonghyeon Lee, Dongha Lee and Hwanjo Yu .....	590
<i>Understanding and Countering Stereotypes: A Computational Approach to the Stereotype Content Model</i> Kathleen C. Fraser, Isar Nejadgholi and Svetlana Kiritchenko .....	600
<i>Structurizing Misinformation Stories via Rationalizing Fact-Checks</i> Shan Jiang and Christo Wilson .....	617

<i>Modeling Language Usage and Listener Engagement in Podcasts</i> Sravana Reddy, Mariya Lazarova, Yongze Yu and Rosie Jones .....	632
<i>Breaking Down the Invisible Wall of Informal Fallacies in Online Discussions</i> Saumya Sahai, Oana Balalau and Roxana Horincar .....	644
<i>SocAoG: Incremental Graph Parsing for Social Relation Inference in Dialogues</i> Liang Qiu, Yuan Liang, Yizhou Zhao, Pan Lu, Baolin Peng, Zhou Yu, Ying Nian Wu and Song-Chun Zhu .....	658
<i>TicketTalk: Toward human-level performance with end-to-end, transaction-based dialog systems</i> Bill Byrne, Karthik Krishnamoorthi, Saravanan Ganesh and Mihir Kale .....	671
<i>Improving Dialog Systems for Negotiation with Personality Modeling</i> Runzhe Yang, Jingxiao Chen and Karthik Narasimhan .....	681
<i>Learning from Perturbations: Diverse and Informative Dialogue Generation with Inverse Adversarial Training</i> Wangchunshu Zhou, Qifei LI and Chenle Li .....	694
<i>Increasing Faithfulness in Knowledge-Grounded Dialogue with Controllable Features</i> Hannah Rashkin, David Reitter, Gaurav Singh Tomar and Dipanjan Das .....	704
<i>CitationIE: Leveraging the Citation Graph for Scientific Information Extraction</i> Vijay Viswanathan, Graham Neubig and Pengfei Liu .....	719
<i>From Discourse to Narrative: Knowledge Projection for Event Relation Extraction</i> Jialong Tang, Hongyu Lin, Meng Liao, Yaojie Lu, Xianpei Han, Le Sun, Weijian Xie and Jin Xu	732
<i>AdvPicker: Effectively Leveraging Unlabeled Data via Adversarial Discriminator for Cross-Lingual NER</i> Weile Chen, Huiqiang Jiang, Qianhui Wu, Börje Karlsson and Yi Guan .....	743
<i>Compare to The Knowledge: Graph Neural Fake News Detection with External Knowledge</i> Linmei Hu, Tianchi Yang, Luhao Zhang, Wanjun Zhong, Duyu Tang, Chuan Shi, Nan Duan and Ming Zhou .....	754
<i>Discontinuous Named Entity Recognition as Maximal Clique Discovery</i> Yucheng Wang, Bowen Yu, Hongsong Zhu, Tingwen Liu, Nan Yu and Limin Sun .....	764
<i>LNN-EL: A Neuro-Symbolic Approach to Short-text Entity Linking</i> Hang Jiang, Sairam Gurajada, Qiuhaio Lu, Sumit Neelam, Lucian Popa, Prithviraj Sen, Yunyao Li and Alexander Gray .....	775
<i>Do Context-Aware Translation Models Pay the Right Attention?</i> Kayo Yin, Patrick Fernandes, Danish Pruthi, Aditi Chaudhary, André F. T. Martins and Graham Neubig .....	788
<i>Adapting High-resource NMT Models to Translate Low-resource Related Languages without Parallel Data</i> Wei-Jen Ko, Ahmed El-Kishky, Adithya Renduchintala, Vishrav Chaudhary, Naman Goyal, Francisco Guzmán, Pascale Fung, Philipp Koehn and Mona Diab .....	802

<i>Bilingual Lexicon Induction via Unsupervised Bitext Construction and Word Alignment</i>	
Haoyue Shi, Luke Zettlemoyer and Sida I. Wang	813
<i>Multilingual Speech Translation from Efficient Finetuning of Pretrained Models</i>	
Xian Li, Changhan Wang, Yun Tang, Chau Tran, Yuqing Tang, Juan Pino, Alexei Baevski, Alexis Conneau and Michael Auli	827
<i>Learning Faithful Representations of Causal Graphs</i>	
Ananth Balashankar and Lakshminarayanan Subramanian	839
<i>What Context Features Can Transformer Language Models Use?</i>	
Joe O’Connor and Jacob Andreas	851
<i>Integrated Directional Gradients: Feature Interaction Attribution for Neural NLP Models</i>	
Sandipan Sikdar, Parantapa Bhattacharya and Kieran Heese	865
<i>DeCLUTR: Deep Contrastive Learning for Unsupervised Textual Representations</i>	
John Giorgi, Osvald Nitski, Bo Wang and Gary Bader	879
<i>XLPT-AMR: Cross-Lingual Pre-Training via Multi-Task Learning for Zero-Shot AMR Parsing and Text Generation</i>	
Dongqin Xu, Junhui Li, Muhua Zhu, Min Zhang and Guodong Zhou	896
<i>Span-based Semantic Parsing for Compositional Generalization</i>	
Jonathan Herzig and Jonathan Berant	908
<i>Compositional Generalization and Natural Language Variation: Can a Semantic Parsing Approach Handle Both?</i>	
Peter Shaw, Ming-Wei Chang, Panupong Pasupat and Kristina Toutanova	922
<i>A Targeted Assessment of Incremental Processing in Neural Language Models and Humans</i>	
Ethan Wilcox, Pranali Vani and Roger Levy	939
<i>The Possible, the Plausible, and the Desirable: Event-Based Modality Detection for Language Processing</i>	
Valentina Pyatkin, Shoval Sadde, Aynat Rubinstein, Paul Portner and Reut Tsarfaty	953
<i>To POS Tag or Not to POS Tag: The Impact of POS Tags on Morphological Learning in Low-Resource Settings</i>	
Sarah Moeller, Ling Liu and Mans Hulden	966
<i>Prosodic segmentation for parsing spoken dialogue</i>	
Elizabeth Nielsen, Mark Steedman and Sharon Goldwater	979
<i>VoxPopuli: A Large-Scale Multilingual Speech Corpus for Representation Learning, Semi-Supervised Learning and Interpretation</i>	
Changhan Wang, Morgane Riviere, Ann Lee, Anne Wu, Chaitanya Talnikar, Daniel Haziza, Mary Williamson, Juan Pino and Emmanuel Dupoux	993
<i>Stereotyping Norwegian Salmon: An Inventory of Pitfalls in Fairness Benchmark Datasets</i>	
Su Lin Blodgett, Gilsinia Lopez, Alexandra Olteanu, Robert Sim and Hanna Wallach	1004
<i>Robust Knowledge Graph Completion with Stacked Convolutions and a Student Re-Ranking Network</i>	
Justin Lovelace, Denis Newman-Griffis, Shikhar Vashishth, Jill Fain Lehman and Carolyn Rosé	1016

<i>A DQN-based Approach to Finding Precise Evidences for Fact Verification</i>	
Hai Wan, Haicheng Chen, Jianfeng Du, Weilin Luo and Rongzhen Ye .....	1030
<i>The Art of Abstinence: Selective Prediction and Error Regularization for Natural Language Processing</i>	
Ji Xin, Raphael Tang, Yaoliang Yu and Jimmy Lin .....	1040
<i>Unsupervised Out-of-Domain Detection via Pre-trained Transformers</i>	
Keyang Xu, Tongzheng Ren, Shikun Zhang, Yihao Feng and Caiming Xiong .....	1052
<i>MATE-KD: Masked Adversarial Text, a Companion to Knowledge Distillation</i>	
Ahmad Rashid, Vasileios Lioutas and Mehdi Rezagholizadeh .....	1062
<i>Selecting Informative Contexts Improves Language Model Fine-tuning</i>	
Richard Antonello, Nicole Beckage, Javier Turek and Alexander Huth .....	1072
<i>Explainable Prediction of Text Complexity: The Missing Preliminaries for Text Simplification</i>	
Cristina Garbacea, Mengtian Guo, Samuel Carton and Qiaozhu Mei .....	1086
<i>Multi-Task Retrieval for Knowledge-Intensive Tasks</i>	
Jean Maillard, Vladimir Karpukhin, Fabio Petroni, Wen-tau Yih, Barlas Oguz, Veselin Stoyanov and Gargi Ghosh .....	1098
<i>When Do You Need Billions of Words of Pretraining Data?</i>	
Yian Zhang, Alex Warstadt, Xiaocheng Li and Samuel R. Bowman .....	1112
<i>Analyzing the Source and Target Contributions to Predictions in Neural Machine Translation</i>	
Elena Voita, Rico Sennrich and Ivan Titov .....	1126
<i>Comparing Test Sets with Item Response Theory</i>	
Clara Vania, Phu Mon Htut, William Huang, Dhara Mungra, Richard Yuanzhe Pang, Jason Phang, Haokun Liu, Kyunghyun Cho and Samuel R. Bowman .....	1141
<i>Uncovering Constraint-Based Behavior in Neural Models via Targeted Fine-Tuning</i>	
Forrest Davis and Marten van Schijndel .....	1159
<i>More Identifiable yet Equally Performant Transformers for Text Classification</i>	
Rishabh Bhardwaj, Navonil Majumder, Soujanya Poria and Eduard Hovy .....	1172
<i>AugNLG: Few-shot Natural Language Generation using Self-trained Data Augmentation</i>	
Xinnuo Xu, Guoyin Wang, Young-Bum Kim and Sungjin Lee .....	1183
<i>Can vectors read minds better than experts? Comparing data augmentation strategies for the automated scoring of children's mindreading ability</i>	
Venelin Kovatchev, Phillip Smith, Mark Lee and Rory Devine .....	1196
<i>A Dataset and Baselines for Multilingual Reply Suggestion</i>	
Mozhi Zhang, Wei Wang, Budhaditya Deb, Guoqing Zheng, Milad Shokouhi and Ahmed Hassan Awadallah .....	1207
<i>What Ingredients Make for an Effective Crowdsourcing Protocol for Difficult NLU Data Collection Tasks?</i>	
Nikita Nangia, Saku Sugawara, Harsh Trivedi, Alex Warstadt, Clara Vania and Samuel R. Bowman	
1221	



<i>Align Voting Behavior with Public Statements for Legislator Representation Learning</i> Xinyi Mou, Zhongyu Wei, Lei Chen, Shangyi Ning, Yancheng He, Changjian Jiang and Xuanjing Huang .....	1236
<i>Measure and Evaluation of Semantic Divergence across Two Languages</i> Syrielle Montariol and Alexandre Allauzen .....	1247
<i>Improving Zero-Shot Translation by Disentangling Positional Information</i> Danni Liu, Jan Niehues, James Cross, Francisco Guzmán and Xian Li .....	1259
<i>Common Sense Beyond English: Evaluating and Improving Multilingual Language Models for Commonsense Reasoning</i> Bill Yuchen Lin, Seyeon Lee, Xiaoyang Qiao and Xiang Ren .....	1274
<i>Attention Calibration for Transformer in Neural Machine Translation</i> Yu Lu, Jiali Zeng, Jiajun Zhang, Shuangzhi Wu and Mu Li .....	1288
<i>Diverse Pretrained Context Encodings Improve Document Translation</i> Domenic Donato, Lei Yu and Chris Dyer .....	1299
<i>Exploiting Language Relatedness for Low Web-Resource Language Model Adaptation: An Indic Languages Study</i> Yash Khemchandani, Sarvesh Mehtani, Vaidehi Patil, Abhijeet Awasthi, Partha Talukdar and Sunita Sarawagi .....	1312
<i>On Finding the K-best Non-projective Dependency Trees</i> Ran Zmigrod, Tim Vieira and Ryan Cotterell .....	1324
<i>Towards Argument Mining for Social Good: A Survey</i> Eva Maria Vecchi, Neele Falk, Iman Jundi and Gabriella Lapesa .....	1338
<i>Automated Generation of Storytelling Vocabulary from Photographs for use in AAC</i> Mauricio Fontana de Vargas and Karyn Moffatt .....	1353
<i>CLIP: A Dataset for Extracting Action Items for Physicians from Hospital Discharge Notes</i> James Mullenbach, Yada Pruksachatkun, Sean Adler, Jennifer Seale, Jordan Swartz, Greg McKelvey, Hui Dai, Yi Yang and David Sontag .....	1365
<i>Assessing Emoji Use in Modern Text Processing Tools</i> Abu Awal Md Shoeb and Gerard de Melo .....	1379
<i>Select, Extract and Generate: Neural Keyphrase Generation with Layer-wise Coverage Attention</i> Wasi Ahmad, Xiao Bai, Soomin Lee and Kai-Wei Chang .....	1389
<i>Factorising Meaning and Form for Intent-Preserving Paraphrasing</i> Tom Hosking and Mirella Lapata .....	1405
<i>AggGen: Ordering and Aggregating while Generating</i> Xinnuo Xu, Ondřej Dušek, Verena Rieser and Ioannis Konstas .....	1419
<i>Reflective Decoding: Beyond Unidirectional Generation with Off-the-Shelf Language Models</i> Peter West, Ximing Lu, Ari Holtzman, Chandra Bhagavatula, Jena D. Hwang and Yejin Choi .....	1435
<i>Towards Table-to-Text Generation with Numerical Reasoning</i> Lya Hulliyiyatus Suadaa, Hidetaka Kamigaito, Kotaro Funakoshi, Manabu Okumura and Hiroya Takamura .....	1451

<i>BACO: A Background Knowledge- and Content-Based Framework for Citing Sentence Generation</i> Yubin Ge, Ly Dinh, Xiaofeng Liu, Jinsong Su, Ziyao Lu, Ante Wang and Jana Diesner . . . . .	1466
<i>Language Model as an Annotator: Exploring DialoGPT for Dialogue Summarization</i> Xiachong Feng, Xiaocheng Feng, Libo Qin, Bing Qin and Ting Liu . . . . .	1479
<i>Challenges in Information-Seeking QA: Unanswerable Questions and Paragraph Retrieval</i> Akari Asai and Eunsol Choi . . . . .	1492
<i>A Gradually Soft Multi-Task and Data-Augmented Approach to Medical Question Understanding</i> Khalil Mrini, Franck Dernoncourt, Seunghyun Yoon, Trung Bui, Walter Chang, Emilia Farcas and Ndapa Nakashole . . . . .	1505
<i>Leveraging Type Descriptions for Zero-shot Named Entity Recognition and Classification</i> Rami Aly, Andreas Vlachos and Ryan McDonald . . . . .	1516
<i>MECT: Multi-Metadata Embedding based Cross-Transformer for Chinese Named Entity Recognition</i> Shuang Wu, Xiaoning Song and Zhenhua Feng . . . . .	1529
<i>Factuality Assessment as Modal Dependency Parsing</i> Jiarui Yao, Haoling Qiu, Jin Zhao, Bonan Min and Nianwen Xue . . . . .	1540
<i>Directed Acyclic Graph Network for Conversational Emotion Recognition</i> Weizhou Shen, Siyue Wu, Yunyi Yang and Xiaojun Quan . . . . .	1551
<i>Improving Formality Style Transfer with Context-Aware Rule Injection</i> Zonghai Yao and hong yu . . . . .	1561
<i>Topic-Driven and Knowledge-Aware Transformer for Dialogue Emotion Detection</i> Lixing Zhu, Gabriele Pergola, Lin Gui, Deyu Zhou and Yulan He . . . . .	1571
<i>Syntopical Graphs for Computational Argumentation Tasks</i> Joe Barrow, Rajiv Jain, Nedim Lipka, Franck Dernoncourt, Vlad Morariu, Varun Manjunatha, Douglas Oard, Philip Resnik and Henning Wachsmuth . . . . .	1583
<i>Stance Detection in COVID-19 Tweets</i> Kyle Glandt, Sarthak Khanal, Yingjie Li, Doina Caragea and Cornelia Caragea . . . . .	1596
<i>Topic-Aware Evidence Reasoning and Stance-Aware Aggregation for Fact Verification</i> Jiasheng Si, Deyu Zhou, Tongzhe Li, Xingyu Shi and Yulan He . . . . .	1612
<i>Changes in European Solidarity Before and During COVID-19: Evidence from a Large Crowd- and Expert-Annotated Twitter Dataset</i> Alexandra Ils, Dan Liu, Daniela Grunow and Steffen Eger . . . . .	1623
<i>Measuring Conversational Uptake: A Case Study on Student-Teacher Interactions</i> Dorottya Demszky, Jing Liu, Zid Mancenido, Julie Cohen, Heather Hill, Dan Jurafsky and Tatsunori Hashimoto . . . . .	1638
<i>A Survey of Code-switching: Linguistic and Social Perspectives for Language Technologies</i> A. Seza Dođruöz, Sunayana Sitaram, Barbara E. Bullock and Almedia Jacqueline Toribio . . .	1654
<i>Learning from the Worst: Dynamically Generated Datasets to Improve Online Hate Detection</i> Bertie Vidgen, Tristan Thrush, Zeerak Waseem and Douwe Kiela . . . . .	1667

<i>InfoSurgeon: Cross-Media Fine-grained Information Consistency Checking for Fake News Detection</i>	
Yi Fung, Christopher Thomas, Revanth Gangi Reddy, Sandeep Polisetty, Heng Ji, Shih-Fu Chang, Kathleen McKeown, Mohit Bansal and Avi Sil .....	1683
<i>I like fish, especially dolphins: Addressing Contradictions in Dialogue Modeling</i>	
Yixin Nie, Mary Williamson, Mohit Bansal, Douwe Kiela and Jason Weston .....	1699
<i>A Sequence-to-Sequence Approach to Dialogue State Tracking</i>	
Yue Feng, Yang Wang and Hang Li .....	1714
<i>Discovering Dialog Structure Graph for Coherent Dialog Generation</i>	
Jun Xu, Zeyang Lei, Haifeng Wang, Zheng-Yu Niu, Hua Wu and Wanxiang Che .....	1726
<i>Dialogue Response Selection with Hierarchical Curriculum Learning</i>	
Yixuan Su, Deng Cai, Qingyu Zhou, Zibo Lin, Simon Baker, Yunbo Cao, Shuming Shi, Nigel Collier and Yan Wang .....	1740
<i>A Joint Model for Dropped Pronoun Recovery and Conversational Discourse Parsing in Chinese Conversational Speech</i>	
Jingxuan Yang, Kerui Xu, Jun Xu, Si Li, Sheng Gao, Jun Guo, Nianwen Xue and Ji-Rong Wen	1752
<i>A Systematic Investigation of KB-Text Embedding Alignment at Scale</i>	
Vardaan Pahuja, Yu Gu, Wenhu Chen, Mehdi Bahrami, Lei Liu, Wei-Peng Chen and Yu Su ..	1764
<i>Named Entity Recognition with Small Strongly Labeled and Large Weakly Labeled Data</i>	
Haoming Jiang, Danqing Zhang, Tianyu Cao, Bing Yin and Tuo Zhao .....	1775
<i>Ultra-Fine Entity Typing with Weak Supervision from a Masked Language Model</i>	
Hongliang Dai, Yangqiu Song and Haixun Wang .....	1790
<i>Improving Named Entity Recognition by External Context Retrieving and Cooperative Learning</i>	
Xinyu Wang, Yong Jiang, Nguyen Bach, Tao Wang, Zhongqiang Huang, Fei Huang and Kewei Tu	1800
<i>Implicit Representations of Meaning in Neural Language Models</i>	
Belinda Z. Li, Maxwell Nye and Jacob Andreas .....	1813
<i>Causal Analysis of Syntactic Agreement Mechanisms in Neural Language Models</i>	
Matthew Finlayson, Aaron Mueller, Sebastian Gehrmann, Stuart Shieber, Tal Linzen and Yonatan Belinkov .....	1828
<i>Bird's Eye: Probing for Linguistic Graph Structures with a Simple Information-Theoretic Approach</i>	
Yifan Hou and Mrinmaya Sachan .....	1844
<i>Knowledgeable or Educated Guess? Revisiting Language Models as Knowledge Bases</i>	
Boxi Cao, Hongyu Lin, Xianpei Han, Le Sun, Lingyong Yan, Meng Liao, Tong Xue and Jin Xu	1860
<i>Poisoning Knowledge Graph Embeddings via Relation Inference Patterns</i>	
Peru Bhardwaj, John Kelleher, Luca Costabello and Declan O'Sullivan .....	1875
<i>Bad Seeds: Evaluating Lexical Methods for Bias Measurement</i>	
Maria Antoniak and David Mimno .....	1889
<i>A Survey of Race, Racism, and Anti-Racism in NLP</i>	
Anjalie Field, Su Lin Blodgett, Zeerak Waseem and Yulia Tsvetkov .....	1905

<i>Intrinsic Bias Metrics Do Not Correlate with Application Bias</i>	
Seraphina Goldfarb-Tarrant, Rebecca Marchant, Ricardo Muñoz Sánchez, Mugdha Pandya and Adam Lopez .....	1926
<i>RedditBias: A Real-World Resource for Bias Evaluation and Debiasing of Conversational Language Models</i>	
Soumya Barikeri, Anne Lauscher, Ivan Vulić and Goran Glavaš .....	1941
<i>Contributions of Transformer Attention Heads in Multi- and Cross-lingual Tasks</i>	
Weicheng Ma, Kai Zhang, Renze Lou, Lili Wang and Soroush Vosoughi .....	1956
<i>Crafting Adversarial Examples for Neural Machine Translation</i>	
Xinze Zhang, Junzhe Zhang, Zhenhua Chen and Kun He .....	1967
<i>UXLA: A Robust Unsupervised Data Augmentation Framework for Zero-Resource Cross-Lingual NLP</i>	
M Saiful Bari, Tasnim Mohiuddin and Shafiq Joty .....	1978
<i>Glancing Transformer for Non-Autoregressive Neural Machine Translation</i>	
Lihua Qian, Hao Zhou, Yu Bao, Mingxuan Wang, Lin Qiu, Weinan Zhang, Yong Yu and Lei Li	1993
<i>Hierarchical Context-aware Network for Dense Video Event Captioning</i>	
Lei Ji, Xianglin Guo, Haoyang Huang and Xilin Chen .....	2004
<i>Control Image Captioning Spatially and Temporally</i>	
Kun Yan, Lei Ji, Huaishao Luo, Ming Zhou, Nan Duan and Shuai Ma .....	2014
<i>Edited Media Understanding Frames: Reasoning About the Intent and Implications of Visual Misinformation</i>	
Jeff Da, Maxwell Forbes, Rowan Zellers, Anthony Zheng, Jena D. Hwang, Antoine Bosselut and Yejin Choi .....	2026
<i>PIGLeT: Language Grounding Through Neuro-Symbolic Interaction in a 3D World</i>	
Rowan Zellers, Ari Holtzman, Matthew Peters, Roozbeh Mottaghi, Aniruddha Kembhavi, Ali Farhadi and Yejin Choi .....	2040
<i>Modeling Fine-Grained Entity Types with Box Embeddings</i>	
Yasumasa Onoe, Michael Boratko, Andrew McCallum and Greg Durrett .....	2051
<i>ChineseBERT: Chinese Pretraining Enhanced by Glyph and Pinyin Information</i>	
zijun sun, Xiaoya Li, Xiaofei Sun, Yuxian Meng, Xiang Ao, Qing He, Fei Wu and Jiwei Li ..	2065
<i>Weight Distillation: Transferring the Knowledge in Neural Network Parameters</i>	
Ye Lin, Yanyang Li, Ziyang Wang, Bei Li, Quan Du, Tong Xiao and Jingbo Zhu .....	2076
<i>Optimizing Deeper Transformers on Small Datasets</i>	
Peng Xu, Dhruv Kumar, Wei Yang, Wenjie Zi, Keyi Tang, Chenyang Huang, Jackie Chi Kit Cheung, Simon J.D. Prince and Yanshuai Cao .....	2089
<i>BERTAC: Enhancing Transformer-based Language Models with Adversarially Pretrained Convolutional Neural Networks</i>	
Jong-Hoon Oh, Ryu Iida, Julien Kloetzer and Kentaro Torisawa .....	2103
<i>COVID-Fact: Fact Extraction and Verification of Real-World Claims on COVID-19 Pandemic</i>	
Arkadiy Saakyan, Tuhin Chakrabarty and Smaranda Muresan .....	2116

<i>Explaining Relationships Between Scientific Documents</i>	
Kelvin Luu, Xinyi Wu, Rik Koncel-Kedziorski, Kyle Lo, Isabel Cachola and Noah A. Smith .	2130
<i>IrEne: Interpretable Energy Prediction for Transformers</i>	
Qingqing Cao, Yash Kumar Lal, Harsh Trivedi, Aruna Balasubramanian and Niranjan Balasubramanian . . . . .	2145
<i>Mitigating Bias in Session-based Cyberbullying Detection: A Non-Compromising Approach</i>	
Lu Cheng, Ahmadreza Mosallanezhad, Yasin Silva, Deborah Hall and Huan Liu . . . . .	2158
<i>PlotCoder: Hierarchical Decoding for Synthesizing Visualization Code in Programmatic Context</i>	
Xinyun Chen, Linyuan Gong, Alvin Cheung and Dawn Song . . . . .	2169
<i>Changing the World by Changing the Data</i>	
Anna Rogers . . . . .	2182
<i>EarlyBERT: Efficient BERT Training via Early-bird Lottery Tickets</i>	
Xiaohan Chen, Yu Cheng, Shuohang Wang, Zhe Gan, Zhangyang Wang and Jingjing Liu . . . . .	2195
<i>On the Effectiveness of Adapter-based Tuning for Pretrained Language Model Adaptation</i>	
Ruidan He, Linlin Liu, Hai Ye, Qingyu Tan, BOSHEG DING, Liying Cheng, Jiawei Low, Lidong Bing and Luo Si . . . . .	2208
<i>Data Augmentation for Text Generation Without Any Augmented Data</i>	
Wei Bi, Huayang Li and Jiacheng Huang . . . . .	2223
<i>Integrating Semantics and Neighborhood Information with Graph-Driven Generative Models for Document Retrieval</i>	
Zijing Ou, Qinliang Su, Jianxing Yu, Bang Liu, Jingwen Wang, Ruihui Zhao, Changyou Chen and Yefeng Zheng . . . . .	2238
<i>SMURF: SeMantic and linguistic UndeRstanding Fusion for Caption Evaluation via Typicality Analysis</i>	
Joshua Feinglass and Yezhou Yang . . . . .	2250
<i>KaggleDBQA: Realistic Evaluation of Text-to-SQL Parsers</i>	
Chia-Hsuan Lee, Oleksandr Polozov and Matthew Richardson . . . . .	2261
<i>QASR: QCRI Aljazeera Speech Resource A Large Scale Annotated Arabic Speech Corpus</i>	
Hamdy Mubarak, Amir Hussein, Shammur Absar Chowdhury and Ahmed Ali . . . . .	2274
<i>An Empirical Study on Hyperparameter Optimization for Fine-Tuning Pre-trained Language Models</i>	
Xueqing Liu and Chi Wang . . . . .	2286
<i>Better than Average: Paired Evaluation of NLP systems</i>	
Maxime Peyrard, Wei Zhao, Steffen Eger and Robert West . . . . .	2301
<i>Chase: A Large-Scale and Pragmatic Chinese Dataset for Cross-Database Context-Dependent Text-to-SQL</i>	
Jiaqi Guo, Ziliang Si, Yu Wang, Qian Liu, Ming Fan, Jian-Guang LOU, Zijiang Yang and Ting Liu	2316
<i>CLINE: Contrastive Learning with Semantic Negative Examples for Natural Language Understanding</i>	
Dong Wang, Ning Ding, Piji Li and Haitao Zheng . . . . .	2332
<i>Tree-Structured Topic Modeling with Nonparametric Neural Variational Inference</i>	
Ziye Chen, Cheng Ding, Zusheng Zhang, Yanghui Rao and Haoran Xie . . . . .	2343

<i>ExCAR: Event Graph Knowledge Enhanced Explainable Causal Reasoning</i>	
Li Du, Xiao Ding, Kai Xiong, Ting Liu and Bing Qin .....	2354
<i>Distributed Representations of Emotion Categories in Emotion Space</i>	
Xiangyu Wang and Chengqing Zong .....	2364
<i>Style is NOT a single variable: Case Studies for Cross-Stylistic Language Understanding</i>	
Dongyeop Kang and Eduard Hovy .....	2376
<i>DynaSent: A Dynamic Benchmark for Sentiment Analysis</i>	
Christopher Potts, Zhengxuan Wu, Atticus Geiger and Douwe Kiela .....	2388
<i>A Hierarchical VAE for Calibrating Attributes while Generating Text using Normalizing Flow</i>	
Bidisha Samanta, Mohit Agrawal and Niloy Ganguly .....	2405
<i>A Unified Generative Framework for Aspect-based Sentiment Analysis</i>	
Hang Yan, Junqi Dai, Tuo Ji, Xipeng Qiu and Zheng Zhang .....	2416
<i>Discovering Dialogue Slots with Weak Supervision</i>	
Vojtěch Hudeček, Ondřej Dušek and Zhou Yu .....	2430
<i>Enhancing the generalization for Intent Classification and Out-of-Domain Detection in SLU</i>	
Yilin Shen, Yen-Chang Hsu, Avik Ray and Hongxia Jin .....	2443
<i>ProtAugment: Intent Detection Meta-Learning through Unsupervised Diverse Paraphrasing</i>	
Thomas Dopierre, Christophe Gravier and Wilfried Logerais .....	2454
<i>Robustness Testing of Language Understanding in Task-Oriented Dialog</i>	
Jiexi Liu, Ryuichi Takanobu, Jiaxin Wen, Dazhen Wan, hongguang li, weiran nie, Cheng LI, Wei Peng and Minlie Huang .....	2467
<i>Comprehensive Study: How the Context Information of Different Granularity Affects Dialogue State Tracking?</i>	
Puhai Yang, Heyan Huang and Xian-Ling Mao .....	2481
<i>OTTERS: One-turn Topic Transitions for Open-Domain Dialogue</i>	
Karin Sevegnani, David M. Howcroft, Ioannis Konstas and Verena Rieser .....	2492
<i>Towards Robustness of Text-to-SQL Models against Synonym Substitution</i>	
Yujian Gan, Xinyun Chen, Qiuping Huang, Matthew Purver, John R. Woodward, Jinxia Xie and Pengsheng Huang .....	2505
<i>KACE: Generating Knowledge Aware Contrastive Explanations for Natural Language Inference</i>	
Qianglong Chen, Feng Ji, Xiangji Zeng, Feng-Lin Li, Ji Zhang, Haiqing Chen and Yin Zhang .....	2516
<i>Self-Guided Contrastive Learning for BERT Sentence Representations</i>	
Taeuk Kim, Kang Min Yoo and Sang-goo Lee .....	2528
<i>LGESQL: Line Graph Enhanced Text-to-SQL Model with Mixed Local and Non-Local Relations</i>	
Ruisheng Cao, Lu Chen, Zhi Chen, Yanbin Zhao, Su Zhu and Kai Yu .....	2541
<i>Multi-stage Pre-training over Simplified Multimodal Pre-training Models</i>	
Tongtong Liu, Fangxiang Feng and Xiaojie WANG .....	2556
<i>Beyond Sentence-Level End-to-End Speech Translation: Context Helps</i>	
Biao Zhang, Ivan Titov, Barry Haddow and Rico Sennrich .....	2566

<i>LayoutLMv2: Multi-modal Pre-training for Visually-rich Document Understanding</i> Yang Xu, Yiheng Xu, Tengchao Lv, Lei Cui, Furu Wei, Guoxin Wang, Yijuan Lu, Dinei Florencio, Cha Zhang, Wanxiang Che, Min Zhang and Lidong Zhou .....	2579
<i>UNIMO: Towards Unified-Modal Understanding and Generation via Cross-Modal Contrastive Learning</i> Wei Li, Can Gao, Guocheng Niu, Xinyan Xiao, Hao Liu, Jiachen Liu, Hua Wu and Haifeng Wang	2592
<i>Missing Modality Imagination Network for Emotion Recognition with Uncertain Missing Modalities</i> Jinming Zhao, Ruichen Li and Qin Jin .....	2608
<i>Stacked Acoustic-and-Textual Encoding: Integrating the Pre-trained Models into Speech Translation Encoders</i> Chen Xu, Bojie Hu, Yanyang Li, Yuhao Zhang, shen huang, Qi Ju, Tong Xiao and Jingbo Zhu	2619
<i>N-ary Constituent Tree Parsing with Recursive Semi-Markov Model</i> Xin Xin, Jinlong Li and Zeqi Tan .....	2631
<i>Automated Concatenation of Embeddings for Structured Prediction</i> Xinyu Wang, Yong Jiang, Nguyen Bach, Tao Wang, Zhongqiang Huang, Fei Huang and Kewei Tu	2643
<i>Multi-View Cross-Lingual Structured Prediction with Minimum Supervision</i> Zechuan Hu, Yong Jiang, Nguyen Bach, Tao Wang, Zhongqiang Huang, Fei Huang and Kewei Tu	2661
<i>The Limitations of Limited Context for Constituency Parsing</i> Yuchen Li and Andrej Risteski .....	2675
<i>Neural Bi-Lexicalized PCFG Induction</i> Songlin Yang, Yanpeng Zhao and Kewei Tu .....	2688
<i>Ruddit: Norms of Offensiveness for English Reddit Comments</i> Rishav Hada, Sohi Sudhir, Pushkar Mishra, Helen Yannakoudakis, Saif M. Mohammad and Ekaterina Shutova .....	2700
<i>Towards Quantifiable Dialogue Coherence Evaluation</i> Zheng Ye, Liucun Lu, Lishan Huang, Liang Lin and Xiaodan Liang .....	2718
<i>Assessing the Representations of Idiomaticity in Vector Models with a Noun Compound Dataset Labeled at Type and Token Levels</i> Marcos Garcia, Tiago Kramer Vieira, Carolina Scarton, Marco Idiart and Aline Villavicencio	2730
<i>Factoring Statutory Reasoning as Language Understanding Challenges</i> Nils Holzenberger and Benjamin Van Durme .....	2742
<i>Evaluating Evaluation Measures for Ordinal Classification and Ordinal Quantification</i> Tetsuya Sakai .....	2759
<i>Interpretable and Low-Resource Entity Matching via Decoupling Feature Learning from Decision Making</i> Zijun Yao, Chengjiang Li, Tiansi Dong, Xin Lv, Jifan Yu, Lei Hou, Juanzi Li, YICHI ZHANG and zelin Dai .....	2770

<i>Locate and Label: A Two-stage Identifier for Nested Named Entity Recognition</i>	
Yongliang Shen, Xinyin Ma, Zeqi Tan, Shuai Zhang, Wen Wang and Weiming Lu . . . . .	2782
<i>Text2Event: Controllable Sequence-to-Structure Generation for End-to-end Event Extraction</i>	
Yaojie Lu, Hongyu Lin, Jin Xu, Xianpei Han, Jialong Tang, Annan Li, Le Sun, Meng Liao and Shaoyi Chen . . . . .	2795
<i>A Large-Scale Chinese Multimodal NER Dataset with Speech Clues</i>	
Dianbo Sui, Zhengkun Tian, Yubo Chen, Kang Liu and Jun Zhao . . . . .	2807
<i>A Neural Transition-based Joint Model for Disease Named Entity Recognition and Normalization</i>	
Zongcheng Ji, Tian Xia, Mei Han and Jing Xiao . . . . .	2819
<i>OntoED: Low-resource Event Detection with Ontology Embedding</i>	
Shumin Deng, Ningyu Zhang, Luoqiu Li, Chen Hui, tou huaixiao, Mosha Chen, Fei Huang and Huajun Chen . . . . .	2828
<i>Self-Training Sampling with Monolingual Data Uncertainty for Neural Machine Translation</i>	
Wenxiang Jiao, Xing Wang, Zhaopeng Tu, Shuming Shi, Michael Lyu and Irwin King . . . . .	2840
<i>Breaking the Corpus Bottleneck for Context-Aware Neural Machine Translation with Cross-Task Pre-training</i>	
Linqing Chen, Junhui Li, Zhengxian Gong, Boxing Chen, Weihua Luo, Min Zhang and Guodong Zhou . . . . .	2851
<i>Guiding Teacher Forcing with Seer Forcing for Neural Machine Translation</i>	
Yang Feng, Shuhao Gu, Dengji Guo, Zhengxin Yang and Chenze Shao . . . . .	2862
<i>Cascade versus Direct Speech Translation: Do the Differences Still Make a Difference?</i>	
Luisa Bentivogli, Mauro Cettolo, Marco Gaido, Alina Karakanta, Alberto Martinelli, Matteo Negri and Marco Turchi . . . . .	2873
<i>Unsupervised Neural Machine Translation for Low-Resource Domains via Meta-Learning</i>	
Cheonbok Park, Yunwon Tae, TaeHee Kim, Soyoung Yang, Mohammad Azam Khan, Lucy Park and Jaegul Choo . . . . .	2888
<i>Lightweight Cross-Lingual Sentence Representation Learning</i>	
Zhuoyuan Mao, Prakhar Gupta, Chenhui Chu, Martin Jaggi and Sadao Kurohashi . . . . .	2902
<i>ERNIE-Doc: A Retrospective Long-Document Modeling Transformer</i>	
SiYu Ding, Junyuan Shang, Shuohuan Wang, Yu Sun, Hao Tian, Hua Wu and Haifeng Wang .	2914
<i>Marginal Utility Diminishes: Exploring the Minimum Knowledge for BERT Knowledge Distillation</i>	
Yuanxin LIU, Fandong Meng, Zheng Lin, Weiping Wang and Jie Zhou . . . . .	2928
<i>Rational LAMOL: A Rationale-based Lifelong Learning Framework</i>	
Kasidis Kanwatchara, Thanapapas Horsuwan, Piyawat Lertvittayakumjorn, Boonserm Kijisirikul and Peerapon Vateekul . . . . .	2942
<i>EnsLM: Ensemble Language Model for Data Diversity by Semantic Clustering</i>	
Zhibin Duan, Hao Zhang, Chaojie Wang, Zhengjue Wang, Bo Chen and Mingyuan Zhou . . . .	2954
<i>LeeBERT: Learned Early Exit for BERT with cross-level optimization</i>	
Wei Zhu . . . . .	2968



<i>Unsupervised Extractive Summarization-Based Representations for Accurate and Explainable Collaborative Filtering</i>	
Reinald Adrian Pugoy and Hung-Yu Kao . . . . .	2981
<i>PLOME: Pre-training with Misspelled Knowledge for Chinese Spelling Correction</i>	
Shulin Liu, Tao Yang, Tianchi Yue, Feng Zhang and Di Wang . . . . .	2991
<i>Competence-based Multimodal Curriculum Learning for Medical Report Generation</i>	
Fenglin Liu, Shen Ge and Xian Wu . . . . .	3001
<i>Learning Syntactic Dense Embedding with Correlation Graph for Automatic Readability Assessment</i>	
Xinying Qiu, Yuan Chen, Hanwu Chen, Jian-Yun Nie, Yuming Shen and Dawei Lu . . . . .	3013
<i>Meta-KD: A Meta Knowledge Distillation Framework for Language Model Compression across Domains</i>	
Haojie Pan, Chengyu Wang, Minghui Qiu, Yichang Zhang, Yaliang Li and Jun Huang . . . . .	3026
<i>A Semantic-based Method for Unsupervised Commonsense Question Answering</i>	
Yilin Niu, Fei Huang, Jiaming Liang, Wenkai Chen, Xiaoyan Zhu and Minlie Huang . . . . .	3037
<i>Explanations for CommonsenseQA: New Dataset and Models</i>	
Shourya Aggarwal, Divyanshu Mandowara, Vishwajeet Agrawal, Dinesh Khandelwal, Parag Singla and Dinesh Garg . . . . .	3050
<i>Few-Shot Question Answering by Pretraining Span Selection</i>	
Ori Ram, Yuval Kirstain, Jonathan Berant, Amir Globerson and Omer Levy . . . . .	3066
<i>UnitedQA: A Hybrid Approach for Open Domain Question Answering</i>	
Hao Cheng, Yelong Shen, Xiaodong Liu, Pengcheng He, Weizhu Chen and Jianfeng Gao . . . . .	3080
<i>Database reasoning over text</i>	
James Thorne, Majid Yazdani, Marzieh Saeidi, Fabrizio Silvestri, Sebastian Riedel and Alon Halevy	3091
<i>Online Learning Meets Machine Translation Evaluation: Finding the Best Systems with the Least Human Effort</i>	
Vânia Mendonça, Ricardo Rei, Luisa Coheur, Alberto Sardinha and Ana Lúcia Santos . . . . .	3105
<i>How Good is Your Tokenizer? On the Monolingual Performance of Multilingual Language Models</i>	
Phillip Rust, Jonas Pfeiffer, Ivan Vulić, Sebastian Ruder and Iryna Gurevych . . . . .	3118
<i>Evaluating morphological typology in zero-shot cross-lingual transfer</i>	
Antonio Martínez-García, Toni Badia and Jeremy Barnes . . . . .	3136
<i>From Machine Translation to Code-Switching: Generating High-Quality Code-Switched Text</i>	
Ishan Tarunesh, Syamantak Kumar and Preethi Jyothi . . . . .	3154
<i>Fast and Accurate Neural Machine Translation with Translation Memory</i>	
Qiuxiang He, Guoping Huang, Qu Cui, Li Li and Lemao Liu . . . . .	3170
<i>Annotating Online Misogyny</i>	
Philine Zeinert, Nanna Inie and Leon Derczynski . . . . .	3181
<i>Few-NERD: A Few-shot Named Entity Recognition Dataset</i>	
Ning Ding, Guangwei Xu, Yulin Chen, Xiaobin Wang, Xu Han, Pengjun Xie, Haitao Zheng and Zhiyuan Liu . . . . .	3198

<i>MultiMET: A Multimodal Dataset for Metaphor Understanding</i>	
Dongyu Zhang, Minghao Zhang, Heting Zhang, Liang Yang and Hongfei LIN	3214
<i>Human-in-the-Loop for Data Collection: a Multi-Target Counter Narrative Dataset to Fight Online Hate Speech</i>	
Margherita Fanton, Helena Bonaldi, Serra Sinem Tekiroğlu and Marco Guerini	3226
<i>Can Generative Pre-trained Language Models Serve As Knowledge Bases for Closed-book QA?</i>	
Cunxiang Wang, Pai Liu and Yue Zhang	3241
<i>Joint Models for Answer Verification in Question Answering Systems</i>	
Zeyu Zhang, Thuy Vu and Alessandro Moschitti	3252
<i>Answering Ambiguous Questions through Generative Evidence Fusion and Round-Trip Prediction</i>	
Yifan Gao, Henghui Zhu, Patrick Ng, Cicero Nogueira dos Santos, Zhiguo Wang, Feng Nan, Dejiao Zhang, Ramesh Nallapati, Andrew O. Arnold and Bing Xiang	3263
<i>TAT-QA: A Question Answering Benchmark on a Hybrid of Tabular and Textual Content in Finance</i>	
Fengbin Zhu, Wenqiang Lei, Youcheng Huang, Chao Wang, Shuo Zhang, Jiancheng Lv, Fuli Feng and Tat-Seng Chua	3277
<i>Modeling Transitions of Focal Entities for Conversational Knowledge Base Question Answering</i>	
Yunshi Lan and Jing Jiang	3288
<i>Evidence-based Factual Error Correction</i>	
James Thorne and Andreas Vlachos	3298
<i>Probabilistic, Structure-Aware Algorithms for Improved Variety, Accuracy, and Coverage of AMR Alignments</i>	
Austin Blodgett and Nathan Schneider	3310
<i>Meta-Learning to Compositionally Generalize</i>	
Henry Conklin, Bailin Wang, Kenny Smith and Ivan Titov	3322
<i>Taming Pre-trained Language Models with N-gram Representations for Low-Resource Domain Adaptation</i>	
Shizhe Diao, Ruijia Xu, Hongjin Su, Yilei Jiang, Yan Song and Tong Zhang	3336
<i>ERICA: Improving Entity and Relation Understanding for Pre-trained Language Models via Contrastive Learning</i>	
Yujia Qin, Yankai Lin, Ryuichi Takanobu, Zhiyuan Liu, Peng Li, Heng Ji, Minlie Huang, Maosong Sun and Jie Zhou	3350
<i>Position Bias Mitigation: A Knowledge-Aware Graph Model for Emotion Cause Extraction</i>	
Hanqi Yan, Lin Gui, Gabriele Pergola and Yulan He	3364
<i>Every Bite Is an Experience: Key Point Analysis of Business Reviews</i>	
Roy Bar-Haim, Lilach Eden, Yoav Kantor, Roni Friedman and Noam Slonim	3376
<i>Structured Sentiment Analysis as Dependency Graph Parsing</i>	
Jeremy Barnes, Robin Kurtz, Stephan Oepen, Lilja Øvrelid and Erik Velldal	3387
<i>Consistency Regularization for Cross-Lingual Fine-Tuning</i>	
Bo Zheng, Li Dong, Shaohan Huang, Wenhui Wang, Zewen Chi, Saksham Singhal, Wanxiang Che, Ting Liu, Xia Song and Furu Wei	3403

<i>Improving Pretrained Cross-Lingual Language Models via Self-Labeled Word Alignment</i>	Zewen Chi, Li Dong, Bo Zheng, Shaohan Huang, Xian-Ling Mao, Heyan Huang and Furu Wei	3418
<i>Rejuvenating Low-Frequency Words: Making the Most of Parallel Data in Non-Autoregressive Translation</i>	Liang Ding, Longyue Wang, Xuebo Liu, Derek F. Wong, Dacheng Tao and Zhaopeng Tu . . . .	3431
<i>G-Transformer for Document-Level Machine Translation</i>	Guangsheng Bao, Yue Zhang, Zhiyang Teng, Boxing Chen and Weihua Luo . . . . .	3442
<i>Prevent the Language Model from being Overconfident in Neural Machine Translation</i>	Mengqi Miao, Fandong Meng, Yijin Liu, Xiao-Hua Zhou and Jie Zhou . . . . .	3456
<i>Towards Emotional Support Dialog Systems</i>	Siyang Liu, Chujie Zheng, Orianna Demasi, Sahand Sabour, Yu Li, Zhou Yu, Yong Jiang and Minlie Huang . . . . .	3469
<i>Novel Slot Detection: A Benchmark for Discovering Unknown Slot Types in the Task-Oriented Dialogue System</i>	Yanan Wu, Zhiyuan Zeng, Keqing He, Hong Xu, Yuanmeng Yan, Huixing Jiang and Weiran Xu	3484
<i>GTM: A Generative Triple-wise Model for Conversational Question Generation</i>	Lei Shen, Fandong Meng, Jinchao Zhang, Yang Feng and Jie Zhou . . . . .	3495
<i>Diversifying Dialog Generation via Adaptive Label Smoothing</i>	Yida Wang, Yinhe Zheng, Yong Jiang and Minlie Huang . . . . .	3507
<i>Out-of-Scope Intent Detection with Self-Supervision and Discriminative Training</i>	Li-Ming Zhan, Haowen Liang, Bo LIU, Lu Fan, Xiao-Ming Wu and Albert Y.S. Lam . . . . .	3521
<i>Document-level Event Extraction via Heterogeneous Graph-based Interaction Model with a Tracker</i>	Runxin Xu, Tianyu Liu, Lei Li and Baobao Chang . . . . .	3533
<i>Nested Named Entity Recognition via Explicitly Excluding the Influence of the Best Path</i>	Yiran Wang, Hiroyuki Shindo, Yuji Matsumoto and Taro Watanabe . . . . .	3547
<i>LearnDA: Learnable Knowledge-Guided Data Augmentation for Event Causality Identification</i>	Xinyu Zuo, Pengfei Cao, Yubo Chen, Kang Liu, Jun Zhao, Weihua Peng and Yuguang Chen .	3558
<i>Revisiting the Negative Data of Distantly Supervised Relation Extraction</i>	Chenhao Xie, Jiaqing Liang, Jingping Liu, Chengsong Huang, Wenhao Huang and Yanghua Xiao	3572
<i>Knowing the No-match: Entity Alignment with Dangling Cases</i>	Zequn Sun, Muhao Chen and Wei Hu . . . . .	3582
<i>Superbizarre Is Not Superb: Derivational Morphology Improves BERT's Interpretation of Complex Words</i>	Valentin Hofmann, Janet Pierrehumbert and Hinrich Schütze . . . . .	3594
<i>BERT is to NLP what AlexNet is to CV: Can Pre-Trained Language Models Identify Analogies?</i>	Asahi Ushio, Luis Espinosa Anke, Steven Schockaert and Jose Camacho-Collados . . . . .	3609
<i>Exploring the Representation of Word Meanings in Context: A Case Study on Homonymy and Synonymy</i>	Marcos Garcia . . . . .	3625

<i>Measuring Fine-Grained Domain Relevance of Terms: A Hierarchical Core-Fringe Approach</i>	
Jie Huang, Kevin Chang, JinJun Xiong and Wen-mei Hwu . . . . .	3641
<i>HERALD: An Annotation Efficient Method to Detect User Disengagement in Social Conversations</i>	
Weixin Liang, Kai-Hui Liang and Zhou Yu . . . . .	3652
<i>Value-Agnostic Conversational Semantic Parsing</i>	
Emmanouil Antonios Platanios, Adam Pauls, Subhro Roy, Yuchen Zhang, Alexander Kyte, Alan Guo, Sam Thomson, Jayant Krishnamurthy, Jason Wolfe, Jacob Andreas and Dan Klein . . . . .	3666
<i>MPC-BERT: A Pre-Trained Language Model for Multi-Party Conversation Understanding</i>	
Jia-Chen Gu, Chongyang Tao, Zhenhua Ling, Can Xu, Xiubo Geng and Daxin Jiang . . . . .	3682
<i>Best of Both Worlds: Making High Accuracy Non-incremental Transformer-based Disfluency Detection Incremental</i>	
Morteza Rohanian and Julian Hough . . . . .	3693
<i>NeuralWOZ: Learning to Collect Task-Oriented Dialogue via Model-Based Simulation</i>	
Sungdong Kim, Minsuk Chang and Sang-Woo Lee . . . . .	3704
<i>CDRNN: Discovering Complex Dynamics in Human Language Processing</i>	
Cory Shain . . . . .	3718
<i>Structural Guidance for Transformer Language Models</i>	
Peng Qian, Tahira Naseem, Roger Levy and Ramón Fernandez Astudillo . . . . .	3735
<i>Surprisal Estimators for Human Reading Times Need Character Models</i>	
Byung-Doh Oh, Christian Clark and William Schuler . . . . .	3746
<i>CogAlign: Learning to Align Textual Neural Representations to Cognitive Language Processing Signals</i>	
Yuqi Ren and Deyi Xiong . . . . .	3758
<i>Self-Attention Networks Can Process Bounded Hierarchical Languages</i>	
Shunyu Yao, Binghui Peng, Christos Papadimitriou and Karthik Narasimhan . . . . .	3770
<i>TextSETTR: Few-Shot Text Style Extraction and Tunable Targeted Restyling</i>	
Parker Riley, Noah Constant, Mandy Guo, Girish Kumar, David Uthus and Zarana Parekh . . .	3786
<i>H-Transformer-ID: Fast One-Dimensional Hierarchical Attention for Sequences</i>	
Zhenhai Zhu and Radu Soricut . . . . .	3801
<i>Making Pre-trained Language Models Better Few-shot Learners</i>	
Tianyu Gao, Adam Fisch and Danqi Chen . . . . .	3816
<i>A Sweet Rabbit Hole by DARCY: Using Honey Pots to Detect Universal Trigger’s Adversarial Attacks</i>	
Thai Le, Noseong Park and Dongwon Lee . . . . .	3831
<i>Towards Propagation Uncertainty: Edge-enhanced Bayesian Graph Convolutional Networks for Rumor Detection</i>	
Lingwei Wei, Dou Hu, Wei Zhou, Zhaojuan Yue and Songlin Hu . . . . .	3845
<i>Label-Specific Dual Graph Neural Network for Multi-Label Text Classification</i>	
Qianwen Ma, Chunyuan Yuan, Wei Zhou and Songlin Hu . . . . .	3855
<i>TAN-NTM: Topic Attention Networks for Neural Topic Modeling</i>	
Madhur Panwar, Shashank Shailabh, Milan Aggarwal and Balaji Krishnamurthy . . . . .	3865

<i>Cross-language Sentence Selection via Data Augmentation and Rationale Training</i> Yanda Chen, Chris Kedzie, Suraj Nair, Petra Galuscakova, Rui Zhang, Douglas Oard and Kathleen McKeown .....	3881
<i>A Neural Model for Joint Document and Snippet Ranking in Question Answering for Large Document Collections</i> Dimitris Pappas and Ion Androutsopoulos .....	3896
<i>W-RST: Towards a Weighted RST-style Discourse Framework</i> Patrick Huber, Wen Xiao and Giuseppe Carenini .....	3908
<i>ABCD: A Graph Framework to Convert Complex Sentences to a Covering Set of Simple Sentences</i> Yanjun Gao, Ting-Hao Huang and Rebecca J. Passonneau .....	3919
<i>Which Linguist Invented the Lightbulb? Presupposition Verification for Question-Answering</i> Najoung Kim, Ellie Pavlick, Burcu Karagol Ayan and Deepak Ramachandran .....	3932
<i>Adversarial Learning for Discourse Rhetorical Structure Parsing</i> Longyin Zhang, Fang Kong and Guodong Zhou .....	3946
<i>Exploring Discourse Structures for Argument Impact Classification</i> Xin Liu, Jiefu Ou, Yangqiu Song and Xin Jiang .....	3958
<i>Point, Disambiguate and Copy: Incorporating Bilingual Dictionaries for Neural Machine Translation</i> Tong Zhang, Long Zhang, Wei Ye, Bo Li, Jinan Sun, Xiaoyu Zhu, Wen Zhao and Shikun Zhang	3970
<i>VECO: Variable and Flexible Cross-lingual Pre-training for Language Understanding and Generation</i> Fuli Luo, Wei Wang, Jiahao Liu, Yijia Liu, Bin Bi, Songfang Huang, Fei Huang and Luo Si ..	3980
<i>A unified approach to sentence segmentation of punctuated text in many languages</i> Rachel Wicks and Matt Post .....	3995
<i>Towards User-Driven Neural Machine Translation</i> Huan Lin, Liang Yao, Baosong Yang, Dayiheng Liu, Haibo Zhang, Weihua Luo, Degen Huang and Jinsong Su .....	4008
<i>End-to-End Lexically Constrained Machine Translation for Morphologically Rich Languages</i> Josef Jon, João Paulo Aires, Dusan Varis and Ondřej Bojar .....	4019
<i>Handling Extreme Class Imbalance in Technical Logbook Datasets</i> Farhad Akhbardeh, Cecilia Ovesdotter Alm, Marcos Zampieri and Travis Desell .....	4034
<i>ILDC for CJPE: Indian Legal Documents Corpus for Court Judgment Prediction and Explanation</i> Vijit Malik, Rishabh Sanjay, Shubham Kumar Nigam, Kripabandhu Ghosh, Shouvik Kumar Guha, Arnab Bhattacharya and Ashutosh Modi .....	4046
<i>Supporting Cognitive and Emotional Empathic Writing of Students</i> Thiemo Wambganss, Christina Niklaus, Matthias Söllner, Siegfried Handschuh and Jan Marco Leimeister .....	4063
<i>Dual Reader-Parser on Hybrid Textual and Tabular Evidence for Open Domain Question Answering</i> Alexander Hanbo Li, Patrick Ng, Peng Xu, Henghui Zhu, Zhiguo Wang and Bing Xiang .....	4078

<i>Generation-Augmented Retrieval for Open-Domain Question Answering</i>	
Yuning Mao, Pengcheng He, Xiaodong Liu, Yelong Shen, Jianfeng Gao, Jiawei Han and Weizhu Chen.....	4089
<i>Check It Again: Progressive Visual Question Answering via Visual Entailment</i>	
Qingyi Si, Zheng Lin, Ming yu Zheng, Peng Fu and Weiping Wang .....	4101
<i>A Mutual Information Maximization Approach for the Spurious Solution Problem in Weakly Supervised Question Answering</i>	
Zhihong Shao, Lifeng Shang, Qun Liu and Minlie Huang.....	4111
<i>Breaking Down Walls of Text: How Can NLP Benefit Consumer Privacy?</i>	
Abhilasha Ravichander, Alan W Black, Thomas Norton, Shomir Wilson and Norman Sadeh .	4125
<i>Supporting Land Reuse of Former Open Pit Mining Sites using Text Classification and Active Learning</i>	
Christopher Schröder, Kim Bürgl, Yves Annanias, Andreas Niekler, Lydia Müller, Daniel Wiegrefe, Christian Bender, Christoph Mengs, Gerik Scheuermann and Gerhard Heyer .....	4141
<i>Reliability Testing for Natural Language Processing Systems</i>	
Samson Tan, Shafiq Joty, Kathy Baxter, Araz Taeihagh, Gregory A. Bennett and Min-Yen Kan	4153
<i>Learning Language and Multimodal Privacy-Preserving Markers of Mood from Mobile Data</i>	
Paul Pu Liang, Terrance Liu, Anna Cai, Michal Muszynski, Ryo Ishii, Nick Allen, Randy Auerbach, David Brent, Ruslan Salakhutdinov and Louis-Philippe Morency.....	4170
<i>Anonymisation Models for Text Data: State of the art, Challenges and Future Directions</i>	
Pierre Lison, Ildikó Pilán, David Sanchez, Montserrat Batet and Lilja Øvrelid .....	4188
<i>End-to-End AMR Coreference Resolution</i>	
Qiankun Fu, Linfeng Song, Wenyu Du and Yue Zhang .....	4204
<i>How is BERT surprised? Layerwise detection of linguistic anomalies</i>	
Bai Li, Zining Zhu, Guillaume Thomas, Yang Xu and Frank Rudzicz .....	4215
<i>Psycholinguistic Tripartite Graph Network for Personality Detection</i>	
Tao Yang, Feifan Yang, Haolan Ouyang and Xiaojun Quan .....	4229
<i>Verb Metaphor Detection via Contextual Relation Learning</i>	
Wei Song, Shuhui Zhou, Ruiji Fu, Ting Liu and Lizhen Liu .....	4240
<i>Improving Speech Translation by Understanding and Learning from the Auxiliary Text Translation Task</i>	
Yun Tang, Juan Pino, Xian Li, Chaghan Wang and Dmitriy Genzel .....	4252
<i>Probing Toxic Content in Large Pre-Trained Language Models</i>	
Nedjma Ousidhoum, Xinran Zhao, Tianqing Fang, Yangqiu Song and Dit-Yan Yeung .....	4262
<i>Societal Biases in Language Generation: Progress and Challenges</i>	
Emily Sheng, Kai-Wei Chang, Prem Natarajan and Nanyun Peng .....	4275
<i>Reservoir Transformers</i>	
Sheng Shen, Alexei Baevski, Ari Morcos, Kurt Keutzer, Michael Auli and Douwe Kiela .....	4294
<i>Subsequence Based Deep Active Learning for Named Entity Recognition</i>	
Puria Radmard, Yassir Fathullah and Aldo Lipani .....	4310

<i>Convolutions and Self-Attention: Re-interpreting Relative Positions in Pre-trained Language Models</i> Tyler Chang, Yifan Xu, Weijian Xu and Zhuowen Tu .....	4322
<i>BinaryBERT: Pushing the Limit of BERT Quantization</i> Haoli Bai, Wei Zhang, Lu Hou, Lifeng Shang, Jin JIN, Xin Jiang, Qun Liu, Michael Lyu and Irwin King .....	4334
<i>Are Pretrained Convolutions Better than Pretrained Transformers?</i> Yi Tay, Mostafa Dehghani, Jai Prakash Gupta, Vamsi Aribandi, Dara Bahri, Zhen Qin and Donald Metzler .....	4349
<i>PairRE: Knowledge Graph Embeddings via Paired Relation Vectors</i> Linlin Chao, Jianshan He, Taifeng Wang and Wei Chu .....	4360
<i>Hierarchy-aware Label Semantics Matching Network for Hierarchical Text Classification</i> Haibin Chen, Qianli Ma, Zhenxi Lin and Jiangyue Yan .....	4370
<i>HiddenCut: Simple Data Augmentation for Natural Language Understanding with Better Generalizability</i> Jiaao Chen, Dinghan Shen, Weizhu Chen and Diyi Yang .....	4380
<i>Neural Stylistic Response Generation with Disentangled Latent Variables</i> Qingfu Zhu, Wei-Nan Zhang, Ting Liu and William Yang Wang .....	4391
<i>Intent Classification and Slot Filling for Privacy Policies</i> Wasi Ahmad, Jianfeng Chi, Tu Le, Thomas Norton, Yuan Tian and Kai-Wei Chang .....	4402
<i>RADDLE: An Evaluation Benchmark and Analysis Platform for Robust Task-oriented Dialog Systems</i> Baolin Peng, Chunyuan Li, Zhu Zhang, Chenguang Zhu, Jinchao Li and Jianfeng Gao .....	4418
<i>Semantic Representation for Dialogue Modeling</i> Xuefeng Bai, Yulong Chen, Linfeng Song and Yue Zhang .....	4430
<i>A Pre-training Strategy for Zero-Resource Response Selection in Knowledge-Grounded Conversations</i> Chongyang Tao, Changyu Chen, Jiazhan Feng, Ji-Rong Wen and Rui Yan .....	4446
<i>Dependency-driven Relation Extraction with Attentive Graph Convolutional Networks</i> Yuanhe Tian, Guimin Chen, Yan Song and Xiang Wan .....	4458
<i>Evaluating Entity Disambiguation and the Role of Popularity in Retrieval-Based NLP</i> Anthony Chen, Pallavi Gudipati, Shayne Longpre, Xiao Ling and Sameer Singh .....	4472
<i>Evaluation Examples are not Equally Informative: How should that change NLP Leaderboards?</i> Pedro Rodriguez, Joe Barrow, Alexander Miserlis Hoyle, John P. Lalor, Robin Jia and Jordan Boyd-Graber .....	4486
<i>Claim Matching Beyond English to Scale Global Fact-Checking</i> Ashkan Kazemi, Kiran Garimella, Devin Gaffney and Scott Hale .....	4504
<i>SemFace: Pre-training Encoder and Decoder with a Semantic Interface for Neural Machine Translation</i> Shuo Ren, Long Zhou, Shujie Liu, Furu Wei, Ming Zhou and Shuai Ma .....	4518
<i>Energy-Based Reranking: Improving Neural Machine Translation Using Energy-Based Models</i> Sumanta Bhattacharyya, Amirmohammad Rooshenas, Subhajit Naskar, Simeng Sun, Mohit Iyyer and Andrew McCallum .....	4528

<i>Syntax-augmented Multilingual BERT for Cross-lingual Transfer</i>	
Wasi Ahmad, Haoran Li, Kai-Wei Chang and Yashar Mehdad . . . . .	4538
<i>How to Adapt Your Pretrained Multilingual Model to 1600 Languages</i>	
Abteen Ebrahimi and Katharina Kann . . . . .	4555
<i>Weakly Supervised Named Entity Tagging with Learnable Logical Rules</i>	
Jiacheng Li, Haibo Ding, Jingbo Shang, Julian McAuley and Zhe Feng . . . . .	4568
<i>Prefix-Tuning: Optimizing Continuous Prompts for Generation</i>	
Xiang Lisa Li and Percy Liang . . . . .	4582
<i>One2Set: Generating Diverse Keyphrases as a Set</i>	
Jiacheng Ye, Tao Gui, Yichao Luo, Yige Xu and Qi Zhang . . . . .	4598
<i>Continuous Language Generative Flow</i>	
Zineng Tang, Shiyue Zhang, Hyounghun Kim and Mohit Bansal . . . . .	4609
<i>TWAG: A Topic-Guided Wikipedia Abstract Generator</i>	
Fangwei Zhu, Shangqing Tu, Jiaxin Shi, Juanzi Li, Lei Hou and Tong Cui . . . . .	4623
<i>ForecastQA: A Question Answering Challenge for Event Forecasting with Temporal Text Data</i>	
Woojeong Jin, Rahul Khanna, Suji Kim, Dong-Ho Lee, Fred Morstatter, Aram Galstyan and Xiang Ren . . . . .	4636
<i>Recursive Tree-Structured Self-Attention for Answer Sentence Selection</i>	
Khalil Mrini, Emilia Farcas and Ndapa Nakashole . . . . .	4651
<i>How Knowledge Graph and Attention Help? A Qualitative Analysis into Bag-level Relation Extraction</i>	
Zikun Hu, Yixin Cao, Lifu Huang and Tat-Seng Chua . . . . .	4662
<i>Trigger is Not Sufficient: Exploiting Frame-aware Knowledge for Implicit Event Argument Extraction</i>	
Kaiwen Wei, Xian Sun, Zequn Zhang, Jingyuan Zhang, Guo Zhi and li jin . . . . .	4672
<i>Element Intervention for Open Relation Extraction</i>	
Fangchao Liu, Lingyong Yan, Hongyu Lin, Xianpei Han and Le Sun . . . . .	4683
<i>AdaTag: Multi-Attribute Value Extraction from Product Profiles with Adaptive Decoding</i>	
Jun Yan, Nasser Zalmout, Yan Liang, Christan Grant, Xiang Ren and Xin Luna Dong . . . . .	4694
<i>CoRI: Collective Relation Integration with Data Augmentation for Open Information Extraction</i>	
Zhengbao Jiang, Jialong Han, BUNYAMIN SISMAN and Xin Luna Dong . . . . .	4706
<i>Benchmarking Scalable Methods for Streaming Cross Document Entity Coreference</i>	
Robert L Logan IV, Andrew McCallum, Sameer Singh and Dan Bikel . . . . .	4717
<i>Search from History and Reason for Future: Two-stage Reasoning on Temporal Knowledge Graphs</i>	
Zixuan Li, Xiaolong Jin, Saiping Guan, Wei Li, Jiafeng Guo, Yuanzhuo Wang and Xueqi Cheng	4732
<i>Employing Argumentation Knowledge Graphs for Neural Argument Generation</i>	
Khalid Al Khatib, Lukas Trautner, Henning Wachsmuth, Yufang Hou and Benno Stein . . . . .	4744
<i>Learning Span-Level Interactions for Aspect Sentiment Triplet Extraction</i>	
Lu Xu, Yew Ken Chia and Lidong Bing . . . . .	4755



<i>On Compositional Generalization of Neural Machine Translation</i>	
Yafu Li, Yongjing Yin, Yulong Chen and Yue Zhang .....	4767
<i>Mask-Align: Self-Supervised Neural Word Alignment</i>	
Chi Chen, Maosong Sun and Yang Liu .....	4781
<i>GWLAN: General Word-Level Autocompletion for Computer-Aided Translation</i>	
Huayang Li, Lemao Liu, Guoping Huang and Shuming Shi .....	4792
<i>De-biasing Distantly Supervised Named Entity Recognition via Causal Intervention</i>	
Wenkai Zhang, Hongyu Lin, Xianpei Han and Le Sun .....	4803
<i>A Span-Based Model for Joint Overlapped and Discontinuous Named Entity Recognition</i>	
Fei Li, ZhiChao Lin, Meishan Zhang and Donghong Ji .....	4814
<i>MLBiNet: A Cross-Sentence Collective Event Detection Network</i>	
Dongfang Lou, Zhilin Liao, Shumin Deng, Ningyu Zhang and Huajun Chen .....	4829
<i>Exploiting Document Structures and Cluster Consistencies for Event Coreference Resolution</i>	
Hieu Minh Tran, Duy Phung and Thien Huu Nguyen .....	4840
<i>StereoRel: Relational Triple Extraction from a Stereoscopic Perspective</i>	
Xuetao Tian, Liping Jing, Lu He and Feng Liu .....	4851
<i>Knowledge-Enriched Event Causality Identification via Latent Structure Induction Networks</i>	
Pengfei Cao, Xinyu Zuo, Yubo Chen, Kang Liu, Jun Zhao, Yuguang Chen and Weihua Peng .	4862
<i>Turn the Combination Lock: Learnable Textual Backdoor Attacks via Word Substitution</i>	
Fanchao Qi, Yuan Yao, Sophia Xu, Zhiyuan Liu and Maosong Sun .....	4873
<i>Parameter-Efficient Transfer Learning with Diff Pruning</i>	
Demi Guo, Alexander Rush and Yoon Kim .....	4884
<i>R2D2: Recursive Transformer based on Differentiable Tree for Interpretable Hierarchical Language Modeling</i>	
Xiang Hu, Haitao Mi, Zujie Wen, Yafang Wang, Yi Su, Jing Zheng and Gerard de Melo .....	4897
<i>Risk Minimization for Zero-shot Sequence Labeling</i>	
Zechuan Hu, Yong Jiang, Nguyen Bach, Tao Wang, Zhongqiang Huang, Fei Huang and Kewei Tu	4909
<i>WARP: Word-level Adversarial ReProgramming</i>	
Karen Hambardzumyan, Hrant Khachatryan and Jonathan May .....	4921
<i>Lexicon Learning for Few Shot Sequence Modeling</i>	
Ekin Akyurek and Jacob Andreas .....	4934
<i>Personalized Transformer for Explainable Recommendation</i>	
Lei Li, Yongfeng Zhang and Li Chen .....	4947
<i>Generating SOAP Notes from Doctor-Patient Conversations Using Modular Summarization Techniques</i>	
Kundan Krishna, Sopan Khosla, Jeffrey Bigham and Zachary C. Lipton .....	4958
<i>Tail-to-Tail Non-Autoregressive Sequence Prediction for Chinese Grammatical Error Correction</i>	
Piji Li and Shuming Shi .....	4973

<i>Early Detection of Sexual Predators in Chats</i> Matthias Vogt, Ulf Leser and Alan Akbik .....	4985
<i>Writing by Memorizing: Hierarchical Retrieval-based Medical Report Generation</i> Xingyi Yang, Muchao Ye, Quanzeng You and Fenglong Ma .....	5000
<i>Concept-Based Label Embedding via Dynamic Routing for Hierarchical Text Classification</i> Xuepeng Wang, Li Zhao, Bing Liu, Tao Chen, Feng Zhang and Di Wang.....	5010
<i>VisualSparta: An Embarrassingly Simple Approach to Large-scale Text-to-Image Search with Weighted Bag-of-words</i> Xiaopeng Lu, Tiancheng Zhao and Kyusong Lee .....	5020
<i>Few-Shot Text Ranking with Meta Adapted Synthetic Weak Supervision</i> Si Sun, Yingzhuo Qian, Zhenghao Liu, Chenyan Xiong, Kaitao Zhang, Jie Bao, Zhiyuan Liu and Paul Bennett.....	5030
<i>Semi-Supervised Text Classification with Balanced Deep Representation Distributions</i> Changchun Li, Ximing Li and Jihong Ouyang .....	5044
<i>Improving Document Representations by Generating Pseudo Query Embeddings for Dense Retrieval</i> Hongyin Tang, Xingwu Sun, Beihong Jin, Jingang Wang, Fuzheng Zhang and Wei Wu .....	5054
<i>ConSERT: A Contrastive Framework for Self-Supervised Sentence Representation Transfer</i> Yuanmeng Yan, Rumei Li, Sirui Wang, Fuzheng Zhang, Wei Wu and Weiran Xu.....	5065
<i>Exploring Dynamic Selection of Branch Expansion Orders for Code Generation</i> Hui Jiang, Chulun Zhou, Fandong Meng, Biao Zhang, Jie Zhou, Degen Huang, Qingqiang Wu and Jinsong Su.....	5076
<i>COINS: Dynamically Generating CONTEXTUALIZED Inference Rules for Narrative Story Completion</i> Debjit Paul and Anette Frank .....	5086
<i>Reasoning over Entity-Action-Location Graph for Procedural Text Understanding</i> Hao Huang, Xiubo Geng, Jian Pei, Guodong Long and Daxin Jiang .....	5100
<i>From Paraphrasing to Semantic Parsing: Unsupervised Semantic Parsing via Synchronous Semantic Decoding</i> Shan Wu, Bo Chen, Chunlei Xin, Xianpei Han, Le Sun, Weipeng Zhang, Jiansong Chen, Fan Yang and Xunliang Cai .....	5110
<i>Pre-training Universal Language Representation</i> Yian Li and Hai Zhao .....	5122
<i>Structural Pre-training for Dialogue Comprehension</i> Zhuosheng Zhang and Hai Zhao .....	5134
<i>AutoTinyBERT: Automatic Hyper-parameter Optimization for Efficient Pre-trained Language Models</i> Yichun Yin, Cheng Chen, Lifeng Shang, Xin Jiang, Xiao Chen and Qun Liu.....	5146
<i>Data Augmentation with Adversarial Training for Cross-Lingual NLI</i> Xin Dong, Yaxin Zhu, Zuohui Fu, Dongkuan Xu and Gerard de Melo.....	5158
<i>Bootstrapped Unsupervised Sentence Representation Learning</i> Yan Zhang, Ruidan He, ZUOZHU LIU, Lidong Bing and Haizhou Li.....	5168

<i>Learning Event Graph Knowledge for Abductive Reasoning</i>	
Li Du, Xiao Ding, Ting Liu and Bing Qin .....	5181
<i>A Cognitive Regularizer for Language Modeling</i>	
Jason Wei, Clara Meister and Ryan Cotterell .....	5191
<i>Lower Perplexity is Not Always Human-Like</i>	
Tatsuki Kuribayashi, Yohei Oseki, Takumi Ito, Ryo Yoshida, Masayuki Asahara and Kentaro Inui	
5203	
<i>Word Sense Disambiguation: Towards Interactive Context Exploitation from Both Word and Sense Perspectives</i>	
Ming Wang and Yinglin Wang .....	5218
<i>A Knowledge-Guided Framework for Frame Identification</i>	
Xuefeng Su, Ru Li, Xiaoli Li, Jeff Z. Pan, Hu Zhang, Qinghua Chai and Xiaoqi Han .....	5230
<i>Obtaining Better Static Word Embeddings Using Contextual Embedding Models</i>	
Prakhar Gupta and Martin Jaggi .....	5241
<i>Meta-Learning with Variational Semantic Memory for Word Sense Disambiguation</i>	
Yingjun Du, Nithin Holla, Xiantong Zhen, Cees Snoek and Ekaterina Shutova .....	5254
<i>LexFit: Lexical Fine-Tuning of Pretrained Language Models</i>	
Ivan Vulić, Edoardo Maria Ponti, Anna Korhonen and Goran Glavaš .....	5269
<i>Text-Free Image-to-Speech Synthesis Using Learned Segmental Units</i>	
Wei-Ning Hsu, David Harwath, Tyler Miller, Christopher Song and James Glass .....	5284
<i>CTFN: Hierarchical Learning for Multimodal Sentiment Analysis Using Coupled-Translation Fusion Network</i>	
Jiajia Tang, Kang Li, Xuanyu Jin, Andrzej Cichocki, Qibin Zhao and Wanzeng Kong .....	5301
<i>Positional Artefacts Propagate Through Masked Language Model Embeddings</i>	
Ziyang Luo, Artur Kulmizev and Xiaoxi Mao .....	5312
<i>Language Model Evaluation Beyond Perplexity</i>	
Clara Meister and Ryan Cotterell .....	5328
<i>Learning to Explain: Generating Stable Explanations Fast</i>	
Xuelin Situ, Ingrid Zukerman, Cecile Paris, Sameen Maruf and Gholamreza Haffari .....	5340
<i>StereoSet: Measuring stereotypical bias in pretrained language models</i>	
Moin Nadeem, Anna Bethke and Siva Reddy .....	5356
<i>Alignment Rationale for Natural Language Inference</i>	
Zhongtao Jiang, Yuanzhe Zhang, Zhao Yang, Jun Zhao and Kang Liu .....	5372
<i>Enabling Lightweight Fine-tuning for Pre-trained Language Model Compression based on Matrix Product Operators</i>	
Peiyu Liu, Ze-Feng Gao, Wayne Xin Zhao, Zhi-Yuan Xie, Zhong-Yi Lu and Ji-Rong Wen ...	5388
<i>On Sample Based Explanation Methods for NLP: Faithfulness, Efficiency and Semantic Evaluation</i>	
Wei Zhang, Ziming Huang, Yada Zhu, Guangan Ye, Xiaodong Cui and Fan Zhang .....	5399

<i>Syntax-Enhanced Pre-trained Model</i>	
Zenan Xu, Daya Guo, Duyu Tang, Qinliang Su, Linjun Shou, Ming Gong, Wanjun Zhong, Xiaojun Quan, Daxin Jiang and Nan Duan .....	5412
<i>Matching Distributions between Model and Data: Cross-domain Knowledge Distillation for Unsupervised Domain Adaptation</i>	
Bo Zhang, Xiaoming Zhang, Yun Liu, Lei Cheng and Zhoujun Li .....	5423
<i>Counterfactual Inference for Text Classification Debiasing</i>	
Chen Qian, Fuli Feng, Lijie Wen, Chunping Ma and Pengjun Xie .....	5434
<i>HieRec: Hierarchical User Interest Modeling for Personalized News Recommendation</i>	
Tao Qi, Fangzhao Wu, Chuhan Wu, Peiru Yang, Yang Yu, Xing Xie and Yongfeng Huang ...	5446
<i>PP-Rec: News Recommendation with Personalized User Interest and Time-aware News Popularity</i>	
Tao Qi, Fangzhao Wu, Chuhan Wu and Yongfeng Huang .....	5457
<i>Article Reranking by Memory-Enhanced Key Sentence Matching for Detecting Previously Fact-Checked Claims</i>	
Qiang Sheng, Juan Cao, Xueyao Zhang, Xirong Li and Lei Zhong .....	5468
<i>Defense against Synonym Substitution-based Adversarial Attacks via Dirichlet Neighborhood Ensemble</i>	
Yi Zhou, Xiaoqing Zheng, Cho-Jui Hsieh, Kai-Wei Chang and Xuanjing Huang .....	5482
<i>Shortformer: Better Language Modeling using Shorter Inputs</i>	
Ofir Press, Noah A. Smith and Mike Lewis .....	5493
<i>BanditMTL: Bandit-based Multi-task Learning for Text Classification</i>	
Yuren Mao, Zekai Wang, Weiwei Liu, Xuemin Lin and Wenbin Hu .....	5506
<i>Unified Interpretation of Softmax Cross-Entropy and Negative Sampling: With Case Study for Knowledge Graph Embedding</i>	
Hidetaka Kamigaito and Katsuhiko Hayashi .....	5517
<i>De-Confounded Variational Encoder-Decoder for Logical Table-to-Text Generation</i>	
Wenqing Chen, Jidong Tian, Yitian Li, Hao He and Yaohui Jin .....	5532
<i>Rethinking Stealthiness of Backdoor Attack against NLP Models</i>	
Wenkai Yang, Yankai Lin, Peng Li, Jie Zhou and Xu Sun .....	5543
<i>Crowdsourcing Learning as Domain Adaptation: A Case Study on Named Entity Recognition</i>	
Xin Zhang, Guangwei Xu, Yueheng Sun, Meishan Zhang and Pengjun Xie .....	5558
<i>Exploring Distantly-Labeled Rationales in Neural Network Models</i>	
Quzhe Huang, Shengqi Zhu, Yansong Feng and Dongyan Zhao .....	5571
<i>Learning to Perturb Word Embeddings for Out-of-distribution QA</i>	
Seanie Lee, Minki Kang, Juho Lee and Sung Ju Hwang .....	5583
<i>Maria: A Visual Experience Powered Conversational Agent</i>	
Zujie Liang, Huang Hu, Can Xu, Chongyang Tao, Xiubo Geng, yining Chen, Fan Liang and Daxin Jiang .....	5596
<i>A Human-machine Collaborative Framework for Evaluating Malevolence in Dialogues</i>	
Yangjun Zhang, Pengjie Ren and Maarten de Rijke .....	5612

<i>Generating Relevant and Coherent Dialogue Responses using Self-Separated Conditional Variational AutoEncoders</i>	
Bin Sun, Shaoxiong Feng, Yiwei Li, Jiamou Liu and Kan Li . . . . .	5624
<i>Learning to Ask Conversational Questions by Optimizing Levenshtein Distance</i>	
Zhongkun Liu, Pengjie Ren, Zhumin CHEN, Zhaochun Ren, Maarten de Rijke and Ming Zhou	5638
<i>DVD: A Diagnostic Dataset for Multi-step Reasoning in Video Grounded Dialogue</i>	
Hung Le, Chinnadhurai Sankar, Seungwhan Moon, Ahmad Beirami, Alborz Geramifard and Satwik Kottur . . . . .	5651
<i>MMGCN: Multimodal Fusion via Deep Graph Convolution Network for Emotion Recognition in Conversation</i>	
Jingwen Hu, Yuchen Liu, Jinming Zhao and Qin Jin . . . . .	5666
<i>DynaEval: Unifying Turn and Dialogue Level Evaluation</i>	
Chen Zhang, Yiming Chen, Luis Fernando D’Haro, Yan Zhang, Thomas Friedrichs, Grandee Lee and Haizhou Li . . . . .	5676
<i>CoSQA: 20,000+ Web Queries for Code Search and Question Answering</i>	
Junjie Huang, Duyu Tang, Linjun Shou, Ming Gong, Ke Xu, Daxin Jiang, Ming Zhou and Nan Duan . . . . .	5690
<i>Rewriter-Evaluator Architecture for Neural Machine Translation</i>	
Yangming Li and Kaisheng Yao . . . . .	5701
<i>Modeling Bilingual Conversational Characteristics for Neural Chat Translation</i>	
Yunlong Liang, Fandong Meng, Yufeng Chen, Jinan Xu and Jie Zhou . . . . .	5711
<i>Importance-based Neuron Allocation for Multilingual Neural Machine Translation</i>	
Wanying Xie, Yang Feng, Shuhao Gu and Dong Yu . . . . .	5725
<i>Transfer Learning for Sequence Generation: from Single-source to Multi-source</i>	
Xuancheng Huang, jingfang xu, Maosong Sun and Yang Liu . . . . .	5738
<i>A Closer Look at Few-Shot Crosslingual Transfer: The Choice of Shots Matters</i>	
Mengjie Zhao, Yi Zhu, Ehsan Shareghi, Ivan Vulić, Roi Reichart, Anna Korhonen and Hinrich Schütze . . . . .	5751
<i>Coreference Reasoning in Machine Reading Comprehension</i>	
Mingzhu Wu, Nafise Sadat Moosavi, Dan Roth and Iryna Gurevych . . . . .	5768
<i>Adapting Unsupervised Syntactic Parsing Methodology for Discourse Dependency Parsing</i>	
Liwen Zhang, Ge Wang, Wenjuan Han and Kewei Tu . . . . .	5782
<i>A Conditional Splitting Framework for Efficient Constituency Parsing</i>	
Thanh-Tung Nguyen, Xuan-Phi Nguyen, Shafiq Joty and Xiaoli Li . . . . .	5795
<i>A Unified Generative Framework for Various NER Subtasks</i>	
Hang Yan, Tao Gui, Junqi Dai, Qipeng Guo, Zheng Zhang and Xipeng Qiu . . . . .	5808
<i>An In-depth Study on Internal Structure of Chinese Words</i>	
Chen Gong, Saihao Huang, Houquan Zhou, Zhenghua Li, Min Zhang, Zhefeng Wang, baoxing Huai and Nicholas Jing Yuan . . . . .	5823

<i>MuDA: A Multilingual Data Augmentation Framework for Low-Resource Cross-Lingual NER</i> Linlin Liu, BOSHEING DING, Lidong Bing, Shafiq Joty, Luo Si and Chunyan Miao . . . . .	5834
<i>Lexicon Enhanced Chinese Sequence Labeling Using BERT Adapter</i> Wei Liu, Xiyang Fu, Yue Zhang and Wenming Xiao . . . . .	5847
<i>Math Word Problem Solving with Explicit Numerical Values</i> Qinzhao Wu, Qi Zhang, Zhongyu Wei and Xuanjing Huang . . . . .	5859
<i>Neural-Symbolic Solver for Math Word Problems with Auxiliary Tasks</i> Jinghui Qin, Xiaodan Liang, Yining Hong, Jianheng Tang and Liang Lin . . . . .	5870
<i>SMedBERT: A Knowledge-Enhanced Pre-trained Language Model with Structured Semantics for Medical Text Mining</i> Taolin Zhang, Zerui Cai, Chengyu Wang, Minghui Qiu, Bite Yang and XIAOFENG HE . . . . .	5882
<i>What is Your Article Based On? Inferring Fine-grained Provenance</i> Yi Zhang, Zachary Ives and Dan Roth . . . . .	5894
<i>Cross-modal Memory Networks for Radiology Report Generation</i> Zhihong Chen, Yaling Shen, Yan Song and Xiang Wan . . . . .	5904
<i>Controversy and Conformity: from Generalized to Personalized Aggressiveness Detection</i> Kamil Kanclerz, Alicja Figas, Marcin Gruza, Tomasz Kajdanowicz, Jan Kocon, Daria Puchalska and Przemyslaw Kazienko . . . . .	5915
<i>Multi-perspective Coherent Reasoning for Helpfulness Prediction of Multimodal Reviews</i> Junhao Liu, Zhen Hai, Min Yang and Lidong Bing . . . . .	5927
<i>Instantaneous Grammatical Error Correction with Shallow Aggressive Decoding</i> Xin Sun, Tao Ge, Furu Wei and Houfeng Wang . . . . .	5937
<i>Automatic ICD Coding via Interactive Shared Representation Networks with Self-distillation Mechanism</i> Tong Zhou, Pengfei Cao, Yubo Chen, Kang Liu, Jun Zhao, Kun Niu, Weifeng Chong and Shengping Liu . . . . .	5948
<i>PHMOSpell: Phonological and Morphological Knowledge Guided Chinese Spelling Check</i> Li Huang, Junjie Li, Weiwei Jiang, Zhiyu Zhang, Minchuan Chen, Shaojun Wang and Jing Xiao . . . . .	5958
<i>Guiding the Growth: Difficulty-Controllable Question Generation through Step-by-Step Rewriting</i> Yi Cheng, Siyao Li, Bang Liu, Ruihui Zhao, Sujian Li, Chenghua Lin and Yefeng Zheng . . . . .	5968
<i>Improving Encoder by Auxiliary Supervision Tasks for Table-to-Text Generation</i> Liang Li, Can Ma, Yinliang Yue and Dayong Hu . . . . .	5979
<i>POS-Constrained Parallel Decoding for Non-autoregressive Generation</i> Kexin Yang, Wenqiang Lei, Dayiheng Liu, Weizhen Qi and Jiancheng Lv . . . . .	5990
<i>Bridging Subword Gaps in Pretrain-Finetune Paradigm for Natural Language Generation</i> Xin Liu, Baosong Yang, Dayiheng Liu, Haibo Zhang, Weihua Luo, Min Zhang, Haiying Zhang and Jinsong Su . . . . .	6001
<i>TGEA: An Error-Annotated Dataset and Benchmark Tasks for Text Generation from Pretrained Language Models</i> Jie He, Bo Peng, Yi Liao, Qun Liu and Deyi Xiong . . . . .	6012

<i>Long-Span Summarization via Local Attention and Content Selection</i> Potsawee Manakul and Mark Gales .....	6026
<i>RepSum: Unsupervised Dialogue Summarization based on Replacement Strategy</i> Xiyan Fu, Yating Zhang, Tianyi Wang, Xiaozhong Liu, Changlong Sun and Zhenglu Yang ...	6042
<i>BASS: Boosting Abstractive Summarization with Unified Semantic Graph</i> Wenhao Wu, Wei Li, Xinyan Xiao, Jiachen Liu, Ziqiang Cao, Sujian Li, Hua Wu and Haifeng Wang	6052
<i>Capturing Relations between Scientific Papers: An Abstractive Model for Related Work Section Generation</i> Xiuying Chen, Hind Alamro, Mingzhe Li, Shen Gao, Xiangliang Zhang, Dongyan Zhao and Rui Yan .....	6068
<i>Focus Attention: Promoting Faithfulness and Diversity in Summarization</i> Rahul Aralikkatte, Shashi Narayan, Joshua Maynez, Sascha Rothe and Ryan McDonald .....	6078
<i>Generating Query Focused Summaries from Query-Free Resources</i> Yumo Xu and Mirella Lapata .....	6096
<i>Robustifying Multi-hop QA through Pseudo-Evidentiality Training</i> Kyungjae Lee, Seung-won Hwang, Sang-eun Han and Dohyeon Lee .....	6110
<i>xMoCo: Cross Momentum Contrastive Learning for Open-Domain Question Answering</i> Nan Yang, Furu Wei, Binxing Jiao, Daxing Jiang and Linjun Yang .....	6120
<i>Learn to Resolve Conversational Dependency: A Consistency Training Framework for Conversational Question Answering</i> Gangwoo Kim, Hyunjae Kim, Jungsoo Park and Jaewoo Kang .....	6130
<i>PhotoChat: A Human-Human Dialogue Dataset With Photo Sharing Behavior For Joint Image-Text Modeling</i> Xiaoxue Zang, Lijuan Liu, Maria Wang, Yang Song, Hao Zhang and Jindong Chen .....	6142
<i>Good for Misconceived Reasons: An Empirical Revisiting on the Need for Visual Context in Multimodal Machine Translation</i> Zhiyong Wu, Lingpeng Kong, Wei Bi, Xiang Li and Ben Kao .....	6153
<i>Attend What You Need: Motion-Appearance Synergistic Networks for Video Question Answering</i> Ahjeong Seo, Gi-Cheon Kang, Joonhan Park and Byoung-Tak Zhang .....	6167
<i>BERTifying the Hidden Markov Model for Multi-Source Weakly Supervised Named Entity Recognition</i> Yinghao Li, Pranav Shetty, Lucas Liu, Chao Zhang and Le Song .....	6178
<i>CIL: Contrastive Instance Learning Framework for Distantly Supervised Relation Extraction</i> Tao Chen, Haizhou Shi, Siliang Tang, Zhigang Chen, Fei Wu and Yueting Zhuang .....	6191
<i>SENT: Sentence-level Distant Relation Extraction via Negative Training</i> Ruotian Ma, Tao Gui, Linyang Li, Qi Zhang, Xuanjing Huang and Yaqian Zhou .....	6201
<i>An End-to-End Progressive Multi-Task Learning Framework for Medical Named Entity Recognition and Normalization</i> Baohang Zhou, Xiangrui Cai, Ying Zhang and Xiaojie Yuan .....	6214

<i>PRGC: Potential Relation and Global Correspondence Based Joint Relational Triple Extraction</i>	
Hengyi Zheng, rui wen, Xi Chen, Yifan Yang, Yunyan Zhang, Ziheng Zhang, Ningyu Zhang, Bin Qin, Xu Ming and Yefeng Zheng .....	6225
<i>Learning from Miscellaneous Other-Class Words for Few-shot Named Entity Recognition</i>	
Meihan Tong, Shuai Wang, Bin Xu, Yixin Cao, Minghui Liu, Lei Hou and Juanzi Li .....	6236
<i>Joint Biomedical Entity and Relation Extraction with Knowledge-Enhanced Collective Inference</i>	
Tuan Lai, Heng Ji, ChengXiang Zhai and Quan Hung Tran .....	6248
<i>Fine-grained Information Extraction from Biomedical Literature based on Knowledge-enriched Abstract Meaning Representation</i>	
Zixuan Zhang, Nikolaus Parulian, Heng Ji, Ahmed Elsayed, Skatje Myers and Martha Palmer	6261
<i>Unleash GPT-2 Power for Event Detection</i>	
Amir Pouran Ben Veyseh, Viet Lai, Franck Dernoncourt and Thien Huu Nguyen .....	6271
<i>CLEVE: Contrastive Pre-training for Event Extraction</i>	
Ziqi Wang, Xiaozhi Wang, Xu Han, Yankai Lin, Lei Hou, Zhiyuan Liu, Peng Li, Juanzi Li and Jie Zhou .....	6283
<i>Document-level Event Extraction via Parallel Prediction Networks</i>	
Hang Yang, Dianbo Sui, Yubo Chen, Kang Liu, Jun Zhao and Taifeng Wang .....	6298
<i>StructuralLM: Structural Pre-training for Form Understanding</i>	
Chenliang Li, Bin Bi, Ming Yan, Wei Wang, Songfang Huang, Fei Huang and Luo Si .....	6309
<i>Dual Graph Convolutional Networks for Aspect-based Sentiment Analysis</i>	
Ruifan Li, Hao Chen, Fangxiang Feng, Zhanyu Ma, Xiaojie WANG and Eduard Hovy .....	6319
<i>Multi-Label Few-Shot Learning for Aspect Category Detection</i>	
Mengting Hu, Shiwang Zhao, Honglei Guo, Chao Xue, Hang Gao, Tiegang Gao, renhong cheng and Zhong Su .....	6330
<i>Argument Pair Extraction via Attention-guided Multi-Layer Multi-Cross Encoding</i>	
Liyang Cheng, Tianyu Wu, Lidong Bing and Luo Si .....	6341
<i>A Neural Transition-based Model for Argumentation Mining</i>	
Jianzhu Bao, Chuang Fan, Jipeng Wu, Yixue Dang, Jiachen Du and Ruifeng Xu .....	6354
<i>Keep It Simple: Unsupervised Simplification of Multi-Paragraph Text</i>	
Philippe Laban, Tobias Schnabel, Paul Bennett and Marti A. Hearst .....	6365
<i>Long Text Generation by Modeling Sentence-Level and Discourse-Level Coherence</i>	
Jian Guan, Xiaoxi Mao, changjie fan, Zitao Liu, Wenbiao Ding and Minlie Huang .....	6379
<i>OpenMEVA: A Benchmark for Evaluating Open-ended Story Generation Metrics</i>	
Jian Guan, Zhixin Zhang, Zhuoer Feng, Zitao Liu, Wenbiao Ding, Xiaoxi Mao, changjie fan and Minlie Huang .....	6394
<i>DYPLOC: Dynamic Planning of Content Using Mixed Language Models for Text Generation</i>	
Xinyu Hua, Ashwin Sreevatsa and Lu Wang .....	6408
<i>Controllable Open-ended Question Generation with A New Question Type Ontology</i>	
Shuyang Cao and Lu Wang .....	6424



<i>BERTGen: Multi-task Generation through BERT</i>	
Faidon Mitzalis, Ozan Caglayan, Pranava Madhyastha and Lucia Specia .....	6440
<i>Selective Knowledge Distillation for Neural Machine Translation</i>	
Fusheng Wang, Jianhao Yan, Fandong Meng and Jie Zhou .....	6456
<i>Measuring and Increasing Context Usage in Context-Aware Machine Translation</i>	
Patrick Fernandes, Kayo Yin, Graham Neubig and André F. T. Martins .....	6467
<i>Beyond Offline Mapping: Learning Cross-lingual Word Embeddings through Context Anchoring</i>	
Aitor Ormazabal, Mikel Artetxe, Aitor Soroa, Gorka Labaka and Eneko Agirre .....	6479
<i>CCMatrix: Mining Billions of High-Quality Parallel Sentences on the Web</i>	
Holger Schwenk, Guillaume Wenzek, Sergey Edunov, Edouard Grave, Armand Joulin and Angela Fan .....	6490
<i>Length-Adaptive Transformer: Train Once with Length Drop, Use Anytime with Search</i>	
Gyuwan Kim and Kyunghyun Cho .....	6501
<i>GhostBERT: Generate More Features with Cheap Operations for BERT</i>	
Zhiqi Huang, Lu Hou, Lifeng Shang, Xin Jiang, Xiao Chen and Qun Liu .....	6512
<i>Super Tickets in Pre-Trained Language Models: From Model Compression to Improving Generalization</i>	
Chen Liang, Simiao Zuo, Minshuo Chen, Haoming Jiang, Xiaodong Liu, Pengcheng He, Tuo Zhao and Weizhu Chen .....	6524
<i>A Novel Estimator of Mutual Information for Learning to Disentangle Textual Representations</i>	
Pierre Colombo, Pablo Piantanida and Chloé Clavel .....	6539
<i>Determinantal Beam Search</i>	
Clara Meister, Martina Forster and Ryan Cotterell .....	6551
<i>Multi-hop Graph Convolutional Network with High-order Chebyshev Approximation for Text Reasoning</i>	
Shuoran Jiang, Qingcai Chen, Xin Liu, Baotian Hu and Lisai Zhang .....	6563
<i>Accelerating Text Communication via Abbreviated Sentence Input</i>	
Jiban Adhikary, Jamie Berger and Keith Vertanen .....	6574
<i>Regression Bugs Are In Your Model! Measuring, Reducing and Analyzing Regressions In NLP Model Updates</i>	
YUQING XIE, Yi-An Lai, Yuanjun Xiong, Yi Zhang and Stefano Soatto .....	6589
<i>Detecting Propaganda Techniques in Memes</i>	
Dimitar Dimitrov, Bishr Bin Ali, Shaden Shaar, Firoj Alam, Fabrizio Silvestri, Hamed Firooz, Preslav Nakov and Giovanni Da San Martino .....	6603
<i>On the Efficacy of Adversarial Data Collection for Question Answering: Results from a Large-Scale Randomized Study</i>	
Divyansh Kaushik, Douwe Kiela, Zachary C. Lipton and Wen-tau Yih .....	6618
<i>Learning Dense Representations of Phrases at Scale</i>	
Jinhyuk Lee, Mujeen Sung, Jaewoo Kang and Danqi Chen .....	6634
<i>End-to-End Training of Neural Retrievers for Open-Domain Question Answering</i>	
Devendra Sachan, Mostofa Patwary, Mohammad Shoeybi, Neel Kant, Wei Ping, William L. Hamilton and Bryan Catanzaro .....	6648

<i>Question Answering Over Temporal Knowledge Graphs</i>	
Apoorv Saxena, Soumen Chakrabarti and Partha Talukdar . . . . .	6663
<i>Language Model Augmented Relevance Score</i>	
Ruibo Liu, Jason Wei and Soroush Vosoughi . . . . .	6677
<i>DExperts: Decoding-Time Controlled Text Generation with Experts and Anti-Experts</i>	
Alisa Liu, Maarten Sap, Ximing Lu, Swabha Swayamdipta, Chandra Bhagavatula, Noah A. Smith and Yejin Choi . . . . .	6691
<i>Polyjuice: Generating Counterfactuals for Explaining, Evaluating, and Improving Models</i>	
Tongshuang Wu, Marco Tulio Ribeiro, Jeffrey Heer and Daniel Weld . . . . .	6707
<i>Metaphor Generation with Conceptual Mappings</i>	
Kevin Stowe, Tuhin Chakrabarty, Nanyun Peng, Smaranda Muresan and Iryna Gurevych . . . . .	6724
<i>Learning Latent Structures for Cross Action Phrase Relations in Wet Lab Protocols</i>	
Chaitanya Kulkarni, Jany Chan, Eric Fosler-Lussier and Raghu Machiraju . . . . .	6737
<i>Multimodal Multi-Speaker Merger &amp; Acquisition Financial Modeling: A New Task, Dataset, and Neural Baselines</i>	
Ramit Sawhney, Mihir Goyal, Prakhar Goel, Puneet Mathur and Rajiv Ratn Shah . . . . .	6751
<i>Mid-Air Hand Gestures for Post-Editing of Machine Translation</i>	
Rashad Albo Jamara, Nico Herbig, Antonio Krüger and Josef van Genabith . . . . .	6763
<i>Inter-GPS: Interpretable Geometry Problem Solving with Formal Language and Symbolic Reasoning</i>	
Pan Lu, Ran Gong, Shibiao Jiang, Liang Qiu, Siyuan Huang, Xiaodan Liang and Song-Chun Zhu	6774
<i>Joint Verification and Reranking for Open Fact Checking Over Tables</i>	
Michael Sejr Schlichtkrull, Vladimir Karpukhin, Barlas Oguz, Mike Lewis, Wen-tau Yih and Sebastian Riedel . . . . .	6787
<i>Evaluation of Thematic Coherence in Microblogs</i>	
Iman Munire Bilal, Bo Wang, Maria Liakata, Rob Procter and Adam Tsakalidis . . . . .	6800
<i>Neural semi-Markov CRF for Monolingual Word Alignment</i>	
Wuwei Lan, Chao Jiang and Wei Xu . . . . .	6815
<i>Privacy at Scale: Introducing the PrivaSeer Corpus of Web Privacy Policies</i>	
Mukund Srinath, Shomir Wilson and C Lee Giles . . . . .	6829
<i>The statistical advantage of automatic NLG metrics at the system level</i>	
Johnny Wei and Robin Jia . . . . .	6840
<i>Are Missing Links Predictable? An Inferential Benchmark for Knowledge Graph Completion</i>	
Yixin Cao, Xiang Ji, Xin Lv, Juanzi Li, Yonggang Wen and Hanwang Zhang . . . . .	6855
<i>ConvoSumm: Conversation Summarization Benchmark and Improved Abstractive Summarization with Argument Mining</i>	
Alexander Fabbri, Faiaz Rahman, Imad Rizvi, Borui Wang, Haoran Li, Yashar Mehdad and Dragomir Radev . . . . .	6866

<i>Improving Factual Consistency of Abstractive Summarization via Question Answering</i>	
Feng Nan, Cicero Nogueira dos Santos, Henghui Zhu, Patrick Ng, Kathleen McKeown, Ramesh Nallapati, Dejiao Zhang, Zhiguo Wang, Andrew O. Arnold and Bing Xiang .....	6881
<i>EmailSum: Abstractive Email Thread Summarization</i>	
Shiyue Zhang, Asli Celikyilmaz, Jianfeng Gao and Mohit Bansal .....	6895
<i>Cross-Lingual Abstractive Summarization with Limited Parallel Resources</i>	
Yu Bai, Yang Gao and Heyan Huang .....	6910
<i>Dissecting Generation Modes for Abstractive Summarization Models via Ablation and Attribution</i>	
Jiacheng Xu and Greg Durrett .....	6925
<i>Learning Prototypical Functions for Physical Artifacts</i>	
Tianyu Jiang and Ellen Riloff .....	6941
<i>Verb Knowledge Injection for Multilingual Event Processing</i>	
Olga Majewska, Ivan Vulić, Goran Glavaš, Edoardo Maria Ponti and Anna Korhonen .....	6952
<i>Dynamic Contextualized Word Embeddings</i>	
Valentin Hofmann, Janet Pierrehumbert and Hinrich Schütze .....	6970
<i>Lexical Semantic Change Discovery</i>	
Sinan Kurtiyigit, Maïke Park, Dominik Schlechtweg, Jonas Kuhn and Sabine Schulte im Walde	6985
<i>The R-U-A-Robot Dataset: Helping Avoid Chatbot Deception by Detecting User Questions About Human or Non-Human Identity</i>	
David Gros, Yu Li and Zhou Yu .....	6999
<i>Using Meta-Knowledge Mined from Identifiers to Improve Intent Recognition in Conversational Systems</i>	
Claudio Pinhanez, Paulo Cavalin, Victor Henrique Alves Ribeiro, Ana Appel, Heloisa Candello, Julio Nogima, Mauro Pichiliani, Melina Guerra, Maira de Baysar, Gabriel Malfatti and Henrique Ferreira	7014
<i>Space Efficient Context Encoding for Non-Task-Oriented Dialogue Generation with Graph Attention Transformer</i>	
Fabian Galetzka, Jewgeni Rose, David Schlangen and Jens Lehmann .....	7028
<i>DialogueCRN: Contextual Reasoning Networks for Emotion Recognition in Conversations</i>	
Dou Hu, Lingwei Wei and Xiaoyong Huai .....	7042
<i>Cross-replication Reliability - An Empirical Approach to Interpreting Inter-rater Reliability</i>	
Ka Wong, Praveen Paritosh and Lora Aroyo .....	7053
<i>TIMEDIAL: Temporal Commonsense Reasoning in Dialog</i>	
Lianhui Qin, Aditya Gupta, Shyam Upadhyay, Luheng He, Yejin Choi and Manaal Faruqui ..	7066
<i>RAW-C: Relatedness of Ambiguous Words in Context (A New Lexical Resource for English)</i>	
Sean Trott and Benjamin Bergen .....	7077
<i>ARBERT &amp; MARBERT: Deep Bidirectional Transformers for Arabic</i>	
Muhammad Abdul-Mageed, AbdelRahim Elmadany and El Moatez Billah Nagoudi .....	7088
<i>Improving Paraphrase Detection with the Adversarial Paraphrasing Task</i>	
Animesh Nigohjkar and John Licato .....	7106

<i>ADEPT: An Adjective-Dependent Plausibility Task</i>	
Ali Emami, Ian Porada, Alexandra Olteanu, Kaheer Suleman, Adam Trischler and Jackie Chi Kit Cheung .....	7117
<i>ReadOnce Transformers: Reusable Representations of Text for Transformers</i>	
Shih-Ting Lin, Ashish Sabharwal and Tushar Khot.....	7129
<i>Conditional Generation of Temporally-ordered Event Sequences</i>	
Shih-Ting Lin, Nathanael Chambers and Greg Durrett.....	7142
<i>Hate Speech Detection Based on Sentiment Knowledge Sharing</i>	
Xianbing Zhou, yang yong, xiaochao fan, Ge Ren, Yunfeng Song, Yufeng Diao, Liang Yang and Hongfei LIN.....	7158
<i>Transition-based Bubble Parsing: Improvements on Coordination Structure Prediction</i>	
Tianze Shi and Lillian Lee.....	7167
<i>SpanNER: Named Entity Re-/Recognition as Span Prediction</i>	
Jinlan Fu, Xuanjing Huang and Pengfei Liu.....	7183
<i>StructFormer: Joint Unsupervised Induction of Dependency and Constituency Structure from Masked Language Modeling</i>	
Yikang Shen, Yi Tay, Che Zheng, Dara Bahri, Donald Metzler and Aaron Courville .....	7196
<i>Language Embeddings for Typology and Cross-lingual Transfer Learning</i>	
Dian Yu, Taiqi He and Kenji Sagae.....	7210
<i>Can Sequence-to-Sequence Models Crack Substitution Ciphers?</i>	
Nada Aldarrab and Jonathan May .....	7226
<i>Beyond Noise: Mitigating the Impact of Fine-grained Semantic Divergences on Neural Machine Translation</i>	
Eleftheria Briakou and Marine Carpuat .....	7236
<i>Discriminative Reranking for Neural Machine Translation</i>	
Ann Lee, Michael Auli and Marc’Aurelio Ranzato.....	7250
<i>Mind Your Outliers! Investigating the Negative Impact of Outliers on Active Learning for Visual Question Answering</i>	
Siddharth Karamcheti, Ranjay Krishna, Li Fei-Fei and Christopher Manning .....	7265
<i>All That’s ‘Human’ Is Not Gold: Evaluating Human Evaluation of Generated Text</i>	
Elizabeth Clark, Tal August, Sofia Serrano, Nikita Haduong, Suchin Gururangan and Noah A. Smith.....	7282
<i>Scientific Credibility of Machine Translation Research: A Meta-Evaluation of 769 Papers</i>	
Benjamin Marie, Atsushi Fujita and Raphael Rubino.....	7297
<i>Neural Machine Translation with Monolingual Translation Memory</i>	
Deng Cai, Yan Wang, Huayang Li, Wai Lam and Lemao Liu.....	7307
<i>Intrinsic Dimensionality Explains the Effectiveness of Language Model Fine-Tuning</i>	
Armen Aghajanyan, Sonal Gupta and Luke Zettlemoyer.....	7319
<i>UnNatural Language Inference</i>	
Koustuv Sinha, Prasanna Parthasarathi, Joelle Pineau and Adina Williams.....	7329

<i>Including Signed Languages in Natural Language Processing</i>	
Kayo Yin, Amit Moryossef, Julie Hochgesang, Yoav Goldberg and Malihe Alikhani . . . . .	7347
<i>Vocabulary Learning via Optimal Transport for Neural Machine Translation</i>	
Jingjing Xu, Hao Zhou, Chun Gan, Zaixiang Zheng and Lei Li . . . . .	7361