

Findings

**Findings of the Association for
Computational Linguistics**

Findings of ACL: EMNLP 2021

November 7–11, 2021

©2021 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-955917-10-0

Message from the Program Chairs

Welcome to the *Findings* of ACL: EMNLP 2021! To continue the success of *Findings* of ACL series, we followed EMNLP 2020's initiative to produce this accompanying volume, for papers that narrowly missed acceptance to the main conference, but were judged to be solid, well-executed research, and worthy of publication. Out of the 3,600 submissions under review for EMNLP 2021, 445 papers were invited to be included in the *Findings*. 26 papers declined the offer, leading to 419 (300 long and 119 short) papers to be published in the *Findings* of ACL: EMNLP 2021. The papers accepted to *Findings* equates to 11.6% of the reviewed papers, giving an aggregate acceptance rate for EMNLP and *Findings* of 34.9%. In addition, we also accepted 5 long papers to the *Findings*, through the first ACL Rolling Review (ARR) pilot at EMNLP 2021. These papers are published in this volume as any other *Findings* paper.

Papers published in *Findings* of ACL count as full publications. They are not assigned a presentation slot in the main conference, but rather are published online in a separate volume in the ACL Anthology. There are a number of motivations for this new publication, from allowing timely work to be published quickly, to being more accepting of solid work, and helping to manage the increasing reviewing burden on the community. To increase the visibility of the *Findings* papers, the authors are requested to submit a 6 (for long papers) or 3 minutes (for short papers) pre-recorded video presentation to be included in the virtual conference, hosted by Underline. Our workshop chairs also helped to pair *Findings* papers with EMNLP 2021 workshops, and as a result, around 100 *Findings* papers will be presented at those workshops.

The reviewing process for *Findings* is largely the same as for the main conference and accordingly we wish to thank all involved in EMNLP 2021 for their efforts, as detailed in the Preface to the Proceedings of EMNLP 2021. We would like to specifically thank:

- The whole **Program Committee** for reviewing the submissions, and in particular, the **Senior Area Chairs** for making recommendation on final paper decisions for *Findings*;
- The **Ethics Committee**, chaired by Margot Mieskes and Chris Potts, for their hard work to ensure that all the conditionally accepted *Findings* papers have addressed the ethical issues appropriately;
- **Past *ACL PCs**, including Trevor Cohn, Yulan He and Yang Liu (EMNLP 2020), and Fei Xia, Wenjie Li, Roberto Navigli (ACL-IJCNLP 2021) for sharing their experience on *Findings* papers;
- **Findings Chairs** Gabriel Stanovsky and Tim Van de Cruys, for their tremendous effort in making the volume of *Findings* of ACL: EMNLP 2021;
- **Workshop Chairs** Parisa Kordjamshidi and Minlie Huang, for connecting *Findings* paper authors with workshop organizers for possible presentations.

We hope that *Findings* will continue to serve as a companion to future conferences, and become an important venue for excellent, widely-read, and highly cited work in NLP.

Xuanjing Huang, Fudan University, China
Lucia Specia, Imperial College London, UK
Scott Wen-tau Yih, Facebook, USA

EMNLP 2021 Program Committee Co-Chairs

Table of Contents

<i>K-PLUG: Knowledge-injected Pre-trained Language Model for Natural Language Understanding and Generation in E-Commerce</i>	
Song Xu, Haoran Li, Peng Yuan, Yujia Wang, Youzheng Wu, Xiaodong He, Ying Liu and Bowen Zhou	1
<i>Extracting Topics with Simultaneous Word Co-occurrence and Semantic Correlation Graphs: Neural Topic Modeling for Short Texts</i>	
Yiming Wang, Ximing Li, Xiaotang Zhou and Jihong Ouyang	18
<i>Self-supervised Contrastive Cross-Modality Representation Learning for Spoken Question Answering</i>	
Chenyu You, Nuo Chen and Yuexian Zou	28
<i>Language Clustering for Multilingual Named Entity Recognition</i>	
Kyle Shaffer	40
<i>Neural News Recommendation with Collaborative News Encoding and Structural User Encoding</i>	
Zhiming Mao, Xingshan Zeng and Kam-Fai Wong	46
<i>Self-Teaching Machines to Read and Comprehend with Large-Scale Multi-Subject Question-Answering Data</i>	
Dian Yu, Kai Sun, Dong Yu and Claire Cardie	56
<i>A Web Scale Entity Extraction System</i>	
Xuanting Cai, Quanbin Ma, Jianyu Liu, Pan Li, Qi Zeng, Zhengkan Yang and Pushkar Tripathi ..	69
<i>Joint Multimedia Event Extraction from Video and Article</i>	
Brian Chen, Xudong Lin, Christopher Thomas, Manling Li, Shoya Yoshida, Lovish Chum, Heng Ji and Shih-Fu Chang	74
<i>Fine-grained Semantic Alignment Network for Weakly Supervised Temporal Language Grounding</i>	
Yuechen Wang, Wengang Zhou and Houqiang Li	89
<i>Factual Consistency Evaluation for Text Summarization via Counterfactual Estimation</i>	
Yuexiang Xie, Fei Sun, Yang Deng, Yaliang Li and Bolin Ding	100
<i>Cross-Modal Retrieval Augmentation for Multi-Modal Classification</i>	
Shir Gur, Natalia Neverova, Chris Stauffer, Ser-Nam Lim, Douwe Kiela and Austin Reiter	111
<i>HiTRANS: A Hierarchical Transformer Network for Nested Named Entity Recognition</i>	
Zhiwei Yang, Jing Ma, Hechang Chen, Yunke Zhang and Yi Chang	124
<i>Improving Embedding-based Large-scale Retrieval via Label Enhancement</i>	
Peiyang Liu, Xi Wang, Sen Wang, Wei Ye, Xiangyu Xi and Shikun Zhang	133
<i>Improving Privacy Guarantee and Efficiency of Latent Dirichlet Allocation Model Training Under Differential Privacy</i>	
Tao Huang and Hong Chen	143
<i>Generating Mammography Reports from Multi-view Mammograms with BERT</i>	
Alexander Yalunin, Elena Sokolova, Ilya Burenko, Alexander Ponomarchuk, Olga Puchkova and Dmitriy Umerenkov	153

<i>Euphemistic Phrase Detection by Masked Language Model</i> Wanzheng Zhu and Suma Bhat	163
<i>Decomposing Complex Questions Makes Multi-Hop QA Easier and More Interpretable</i> Ruiliu Fu, Han Wang, Xuejun Zhang, Jun Zhou and Yonghong Yan	169
<i>Segmenting Natural Language Sentences via Lexical Unit Analysis</i> Yangming Li, Lemao Liu and Shuming Shi	181
<i>Dense Hierarchical Retrieval for Open-domain Question Answering</i> Ye Liu, Kazuma Hashimoto, Yingbo Zhou, Semih Yavuz, Caiming Xiong and Philip Yu	188
<i>Visually Grounded Concept Composition</i> Bowen Zhang, Hexiang Hu, Linlu Qiu, Peter Shaw and Fei Sha	201
<i>Compositional Networks Enable Systematic Generalization for Grounded Language Understanding</i> Yen-Ling Kuo, Boris Katz and Andrei Barbu	216
<i>An Unsupervised Method for Building Sentence Simplification Corpora in Multiple Languages</i> Xinyu Lu, Jipeng Qiang, Yun Li, Yunhao Yuan and Yi Zhu	227
<i>WhiteningBERT: An Easy Unsupervised Sentence Embedding Approach</i> Junjie Huang, Duyu Tang, Wanjun Zhong, Shuai Lu, Linjun Shou, Ming Gong, Daxin Jiang and Nan Duan	238
<i>TWEETSUMM - A Dialog Summarization Dataset for Customer Service</i> Guy Feigenblat, Chulaka Gunasekara, Benjamin Sznajder, Sachindra Joshi, David Konopnicki and Ranit Aharonov	245
<i>Discourse-Based Sentence Splitting</i> Liam Cripwell, Joël Legrand and Claire Gardent	261
<i>Multi-Task Dense Retrieval via Model Uncertainty Fusion for Open-Domain Question Answering</i> Minghan Li, Ming Li, Kun Xiong and Jimmy Lin	274
<i>Mining the Cause of Political Decision-Making from Social Media: A Case Study of COVID-19 Policies across the US States</i> Zhijing Jin, Zeyu Peng, Tejas Vaidhya, Bernhard Schoelkopf and Rada Mihalcea	288
<i>Self-Attention Graph Residual Convolutional Networks for Event Detection with dependency relations</i> Anan Liu, Ning Xu and Haozhe Liu	302
<i>Mixup Decoding for Diverse Machine Translation</i> Jicheng Li, Pengzhi Gao, Xuanfu Wu, Yang Feng, Zhongjun He, Hua Wu and Haifeng Wang ..	312
<i>An Alignment-Agnostic Model for Chinese Text Error Correction</i> Liyang Zheng, Yue Deng, Weishun Song, Liang Xu and Jing Xiao	321
<i>Reasoning Visual Dialog with Sparse Graph Learning and Knowledge Transfer</i> Gi-Cheon Kang, Junseok Park, Hwaran Lee, Byoung-Tak Zhang and Jin-Hwa Kim	327
<i>Exploring Sentence Community for Document-Level Event Extraction</i> Yusheng Huang and Weijia Jia	340

<i>A Model of Cross-Lingual Knowledge-Grounded Response Generation for Open-Domain Dialogue Systems</i>	
San Kim, Jin Yea Jang, Minyoung Jung and Saim Shin	352
<i>WHOSE Heritage: Classification of UNESCO World Heritage Statements of "Outstanding Universal Value" with Soft Labels</i>	
Nan Bai, Renqian Luo, Pirouz Nourian and Ana Pereira Roders	366
<i>P-INT: A Path-based Interaction Model for Few-shot Knowledge Graph Completion</i>	
Jingwen Xu, Jing Zhang, Xirui Ke, Yuxiao Dong, Hong Chen, Cuiping Li and Yongbin Liu ...	385
<i>Cartography Active Learning</i>	
Mike Zhang and Barbara Plank	395
<i>Beyond Reptile: Meta-Learned Dot-Product Maximization between Gradients for Improved Single-Task Regularization</i>	
Akhil Kedia, Sai Chetan Chinthakindi and Wonho Ryu	407
<i>GooAQ: Open Question Answering with Diverse Answer Types</i>	
Daniel Khashabi, Amos Ng, Tushar Khot, Ashish Sabharwal, Hannaneh Hajishirzi and Chris Callison-Burch	421
<i>Attention Weights in Transformer NMT Fail Aligning Words Between Sequences but Largely Explain Model Predictions</i>	
Javier Ferrando and Marta R. Costa-jussà	434
<i>BFClass: A Backdoor-free Text Classification Framework</i>	
Zichao Li, Dheeraj Mekala, Chengyu Dong and Jingbo Shang	444
<i>Multilingual Chart-based Constituency Parse Extraction from Pre-trained Language Models</i>	
Taeuk Kim, Bowen Li and Sang-goo Lee	454
<i>Hyperbolic Geometry is Not Necessary: Lightweight Euclidean-Based Models for Low-Dimensional Knowledge Graph Embeddings</i>	
Kai Wang, Yu Liu, Dan Lin and Michael Sheng	464
<i>CascadeBERT: Accelerating Inference of Pre-trained Language Models via Calibrated Complete Models Cascade</i>	
Lei Li, Yankai Lin, Deli Chen, Shuhuai Ren, Peng Li, Jie Zhou and Xu Sun	475
<i>Semi-supervised Relation Extraction via Incremental Meta Self-Training</i>	
Xuming Hu, Chenwei Zhang, Fukun Ma, Chenyao Liu, Lijie Wen and Philip S. Yu	487
<i>Keyphrase Generation with Fine-Grained Evaluation-Guided Reinforcement Learning</i>	
Yichao Luo, Yige Xu, Jiacheng Ye, Xipeng Qiu and Qi Zhang	497
<i>Improving Knowledge Graph Embedding Using Affine Transformations of Entities Corresponding to Each Relation</i>	
Jinfa Yang, Yongjie Shi, Xin Tong, Robin Wang, Taiyan Chen and Xianghua Ying	508
<i>Using Question Answering Rewards to Improve Abstractive Summarization</i>	
Chulaka Gunasekara, Guy Feigenblat, Benjamin Sznajder, Ranit Aharonov and Sachindra Joshi	518
<i>Effect Generation Based on Causal Reasoning</i>	
Feiteng Mu, Wenjie Li and Zhipeng Xie	527

<i>Distilling Word Meaning in Context from Pre-trained Language Models</i> Yuki Arase and Tomoyuki Kajiwara	534
<i>Unseen Entity Handling in Complex Question Answering over Knowledge Base via Language Generation</i> Xin Huang, Jung-Jae Kim and Bowei Zou	547
<i>Bidirectional Hierarchical Attention Networks based on Document-level Context for Emotion Cause Extraction</i> Guimin Hu, Guangming Lu and Yi Zhao	558
<i>Distantly Supervised Relation Extraction in Federated Settings</i> Dianbo Sui, Yubo Chen, Kang Liu and Jun Zhao	569
<i>Casting the Same Sentiment Classification Problem</i> Erik Körner, Ahmad Dawar Hakimi, Gerhard Heyer and Martin Potthast	584
<i>Detecting Compositionally Out-of-Distribution Examples in Semantic Parsing</i> Denis Lukovnikov, Sina Daubener and Asja Fischer	591
<i>Saliency-based Multi-View Mixed Language Training for Zero-shot Cross-lingual Classification</i> Siyu Lai, Hui Huang, Dong Jing, Yufeng Chen, Jinan Xu and Jian Liu	599
<i>Fighting the COVID-19 Infodemic: Modeling the Perspective of Journalists, Fact-Checkers, Social Media Platforms, Policy Makers, and the Society</i> Firoj Alam, Shaden Shaar, Fahim Dalvi, Hassan Sajjad, Alex Nikolov, Hamdy Mubarak, Giovanni Da San Martino, Ahmed Abdelali, Nadir Durrani, Kareem Darwish, Abdulaziz Al-Homaid, Wajdi Zaghoulani, Tommaso Caselli, Gijs Danoe, Friso Stolk, Britt Bruntink and Preslav Nakov	611
<i>FANATIC: FASt Noise-Aware TopIc Clustering</i> Ari Silburt, Anja Subasic, Evan Thompson, Carmeline Dsilva and Tarec Fares	650
<i>Stream-level Latency Evaluation for Simultaneous Machine Translation</i> Javier Iranzo-Sánchez, Jorge Civera Saiz and Alfons Juan	664
<i>TSDAE: Using Transformer-based Sequential Denoising Auto-Encoder for Unsupervised Sentence Embedding Learning</i> Kexin Wang, Nils Reimers and Iryna Gurevych	671
<i>How Suitable Are Subword Segmentation Strategies for Translating Non-Concatenative Morphology?</i> Chantal Amrhein and Rico Sennrich	689
<i>Rethinking Why Intermediate-Task Fine-Tuning Works</i> Ting-Yun Chang and Chi-Jen Lu	706
<i>Learn Continually, Generalize Rapidly: Lifelong Knowledge Accumulation for Few-shot Learning</i> Xisen Jin, Bill Yuchen Lin, Mohammad Rostami and Xiang Ren	714
<i>Efficient Test Time Adapter Ensembling for Low-resource Language Varieties</i> Xinyi Wang, Yulia Tsvetkov, Sebastian Ruder and Graham Neubig	730
<i>An Analysis of Euclidean vs. Graph-Based Framing for Bilingual Lexicon Induction from Word Embedding Spaces</i> Kelly Marchisio, Youngser Park, Ali Saad-Eldin, Anton Alyakin, Kevin Duh, Carey Priebe and Philipp Koehn	738

<i>How to Select One Among All ? An Empirical Study Towards the Robustness of Knowledge Distillation in Natural Language Understanding</i>	
Tianda Li, Ahmad Rashid, Aref Jafari, Pranav Sharma, Ali Ghodsi and Mehdi Rezagholizadeh	750
<i>Recommend for a Reason: Unlocking the Power of Unsupervised Aspect-Sentiment Co-Extraction</i>	
Zeyu Li, Wei Cheng, Reema Kshetramade, John Houser, Haifeng Chen and Wei Wang	763
<i>Learning Hard Retrieval Decoder Attention for Transformers</i>	
Hongfei Xu, Qihui Liu, Josef van Genabith and Deyi Xiong	779
<i>Recall and Learn: A Memory-augmented Solver for Math Word Problems</i>	
Shifeng Huang, Jiawei Wang, Jiao Xu, Da Cao and Ming Yang	786
<i>An Uncertainty-Aware Encoder for Aspect Detection</i>	
Thi-Nhung Nguyen, Kiem-Hieu Nguyen, Young-In Song and Tuan-Dung Cao	797
<i>Improving Empathetic Response Generation by Recognizing Emotion Cause in Conversations</i>	
Jun Gao, Yuhan Liu, Haolin Deng, Wei Wang, Yu Cao, Jiachen Du and Ruifeng Xu	807
<i>Probing Across Time: What Does RoBERTa Know and When?</i>	
Zeyu Liu, Yizhong Wang, Jungo Kasai, Hannaneh Hajishirzi and Noah A. Smith	820
<i>Knowledge-Guided Paraphrase Identification</i>	
Haoyu Wang, Fenglong Ma, Yaqing Wang and Jing Gao	843
<i>R2-D2: A Modular Baseline for Open-Domain Question Answering</i>	
Martin Fajcik, Martin Docekal, Karel Ondrej and Pavel Smrz	854
<i>What Does Your Smile Mean? Jointly Detecting Multi-Modal Sarcasm and Sentiment Using Quantum Probability</i>	
Yaochen Liu, Yazhou Zhang, Qiuchi Li, Benyou Wang and Dawei Song	871
<i>Discovering Representation Sprachbund For Multilingual Pre-Training</i>	
Yimin Fan, Yaobo Liang, Alexandre Muzio, Hany Hassan, Houqiang Li, Ming Zhou and Nan Duan	881
<i>Plan-then-Generate: Controlled Data-to-Text Generation via Planning</i>	
Yixuan Su, David Vandyke, Sihui Wang, Yimai Fang and Nigel Collier	895
<i>Few-Shot Table-to-Text Generation with Prototype Memory</i>	
Yixuan Su, Zaiqiao Meng, Simon Baker and Nigel Collier	910
<i>Leveraging Word-Formation Knowledge for Chinese Word Sense Disambiguation</i>	
Hua Zheng, Lei Li, Damai Dai, Deli Chen, Tianyu Liu, Xu Sun and Yang Liu	918
<i>Exploiting Curriculum Learning in Unsupervised Neural Machine Translation</i>	
Jinliang Lu and Jiajun Zhang	924
<i>Robust Fragment-Based Framework for Cross-lingual Sentence Retrieval</i>	
Nattapol Trijakwanich, Peerat Limkonchotiwat, Raheem Sarwar, Wannaphong Phatthiyaphaibun, Ekapol Chuangsuwanich and Sarana Nutanong	935
<i>Towards Improving Adversarial Training of NLP Models</i>	
Jin Yong Yoo and Yanjun Qi	945

<i>To Protect and To Serve? Analyzing Entity-Centric Framing of Police Violence</i> Caleb Ziems and Diyi Yang	957
<i>Calibrate your listeners! Robust communication-based training for pragmatic speakers</i> Rose Wang, Julia White, Jesse Mu and Noah Goodman	977
<i>When Retriever-Reader Meets Scenario-Based Multiple-Choice Questions</i> ZiXian Huang, Ao Wu, Yulin Shen, Gong Cheng and Yuzhong Qu	985
<i>Structured abbreviation expansion in context</i> Kyle Gorman, Christo Kirov, Brian Roark and Richard Sproat	995
<i>Task-adaptive Pre-training and Self-training are Complementary for Natural Language Understanding</i> Shiyang Li, Semih Yavuz, Wenhua Chen and Xifeng Yan	1006
<i>CNNBiF: CNN-based Bigram Features for Named Entity Recognition</i> Chul Sung, Vaibhava Goel, Etienne Marcheret, Steven Rennie and David Nahamoo	1016
<i>Compositional Generalization via Semantic Tagging</i> Hao Zheng and Mirella Lapata	1022
<i>Towards Document-Level Paraphrase Generation with Sentence Rewriting and Reordering</i> Zhe Lin, Yitao Cai and Xiaojun Wan	1033
<i>Exploring Decomposition for Table-based Fact Verification</i> Xiaoyu Yang and Xiaodan Zhu	1045
<i>Diversity and Consistency: Exploring Visual Question-Answer Pair Generation</i> Sen Yang, Qingyu Zhou, Dawei Feng, Yang Liu, Chao Li, Yunbo Cao and Dongsheng Li	1053
<i>Entity-level Cross-modal Learning Improves Multi-modal Machine Translation</i> Xin Huang, Jiajun Zhang and Chengqing Zong	1067
<i>Learning to Ground Visual Objects for Visual Dialog</i> Feilong Chen, Xiuyi Chen, Can Xu and Daxin Jiang	1081
<i>KERS: A Knowledge-Enhanced Framework for Recommendation Dialog Systems with Multiple Subgoals</i> Jun Zhang, Yan Yang, Chencai Chen, Liang He and Zhou Yu	1092
<i>Less Is More: Domain Adaptation with Lottery Ticket for Reading Comprehension</i> Haichao Zhu, Zekun Wang, Heng Zhang, Ming Liu, Sendong Zhao and Bing Qin	1102
<i>Effectiveness of Pre-training for Few-shot Intent Classification</i> Haode Zhang, Yuwei Zhang, Li-Ming Zhan, Jiabin Chen, Guangyuan Shi, Xiao-Ming Wu and Albert Y.S. Lam	1114
<i>Improving Abstractive Dialogue Summarization with Hierarchical Pretraining and Topic Segment</i> Mengnan Qi, Hao Liu, YuZhuo Fu and Ting Liu	1121
<i>Learning to Answer Psychological Questionnaire for Personality Detection</i> Feifan Yang, Tao Yang, Xiaojun Quan and Qinliang Su	1131
<i>Exploiting Reasoning Chains for Multi-hop Science Question Answering</i> Weiwen Xu, Yang Deng, Huihui Zhang, Deng Cai and Wai Lam	1143

<i>Winnowing Knowledge for Multi-choice Question Answering</i>	
Yequi Li, Bowei Zou, Zhifeng Li, Ai Ti Aw, Yu Hong and Qiaoming Zhu	1157
<i>Neural Media Bias Detection Using Distant Supervision With BABE - Bias Annotations By Experts</i>	
Timo Spinde, Manuel Plank, Jan-David Krieger, Terry Ruas, Bela Gipp and Akiko Aizawa ..	1166
<i>Learning and Evaluating a Differentially Private Pre-trained Language Model</i>	
Shlomo Hoory, Amir Feder, Avichai Tendler, Sofia Erell, Alon Peled-Cohen, Itay Laish, Hootan Nakhost, Uri Stemmer, Ayelet Benjamini, Avinatan Hassidim and Yossi Matias	1178
<i>Simulated Chats for Building Dialog Systems: Learning to Generate Conversations from Instructions</i>	
Biswesh Mohapatra, Gaurav Pandey, Danish Contractor and Sachindra Joshi	1190
<i>Past, Present, and Future: Conversational Emotion Recognition through Structural Modeling of Psychological Knowledge</i>	
Jiangnan Li, Zheng Lin, Peng Fu and Weiping Wang	1204
<i>An unsupervised framework for tracing textual sources of moral change</i>	
Aida Ramezani, Zining Zhu, Frank Rudzicz and Yang Xu	1215
<i>Topic-Aware Contrastive Learning for Abstractive Dialogue Summarization</i>	
Junpeng Liu, Yanyan Zou, Hainan Zhang, Hongshen Chen, Zhuoye Ding, Caixia Yuan and Xiaojie Wang	1229
<i>TWT: Table with Written Text for Controlled Data-to-Text Generation</i>	
Tongliang Li, Lei Fang, Jian-Guang Lou and Zhoujun Li	1244
<i>ArabicTransformer: Efficient Large Arabic Language Model with Funnel Transformer and ELECTRA Objective</i>	
Sultan Alrowili and Vijay Shanker	1255
<i>Which is Making the Contribution: Modulating Unimodal and Cross-modal Dynamics for Multimodal Sentiment Analysis</i>	
Ying Zeng, Sijie Mai and Haifeng Hu	1262
<i>CVAE-based Re-anchoring for Implicit Discourse Relation Classification</i>	
Zujun Dou, Yu Hong, Yu Sun and Guodong Zhou	1275
<i>Combining Curriculum Learning and Knowledge Distillation for Dialogue Generation</i>	
Qingqing Zhu, Xiuying Chen, Pengfei Wu, JunFei Liu and Dongyan Zhao	1284
<i>Improving End-to-End Task-Oriented Dialog System with A Simple Auxiliary Task</i>	
Yohan Lee	1296
<i>EDTC: A Corpus for Discourse-Level Topic Chain Parsing</i>	
Longyin Zhang, Xin Tan, Fang Kong and Guodong Zhou	1304
<i>Multilingual Neural Machine Translation: Can Linguistic Hierarchies Help?</i>	
Fahimeh Saleh, Wray Buntine, Gholamreza Haffari and Lan Du	1313
<i>Self Question-answering: Aspect-based Sentiment Analysis by Role Flipped Machine Reading Comprehension</i>	
Guoxin Yu, Jiwei Li, Ling Luo, Yuxian Meng, Xiang Ao and Qing He	1331
<i>Generalization in Text-based Games via Hierarchical Reinforcement Learning</i>	
Yunqiu Xu, Meng Fang, Ling Chen, Yali Du and Chengqi Zhang	1343

<i>A Finer-grain Universal Dialogue Semantic Structures based Model For Abstractive Dialogue Summarization</i>	
Yuejie Lei, Fujia Zheng, Yuanmeng Yan, Keqing He and Weiran Xu	1354
<i>Constructing contrastive samples via summarization for text classification with limited annotations</i>	
Yangkai Du, Tengfei Ma, Lingfei Wu, Fangli Xu, Xuhong Zhang, Bo Long and Shouling Ji ..	1365
<i>End-to-end Neural Information Status Classification</i>	
Yufang Hou	1377
<i>EventKE: Event-Enhanced Knowledge Graph Embedding</i>	
Zixuan Zhang, Hongwei Wang, Han Zhao, Hanghang Tong and Heng Ji	1389
<i>Modeling Concentrated Cross-Attention for Neural Machine Translation with Gaussian Mixture Model</i>	
Shaolei Zhang and Yang Feng	1401
<i>Inconsistency Matters: A Knowledge-guided Dual-inconsistency Network for Multi-modal Rumor Detection</i>	
Mengzhu Sun, Xi Zhang, Jianqiang Ma and Yazheng Liu	1412
<i>EfficientBERT: Progressively Searching Multilayer Perceptron via Warm-up Knowledge Distillation</i>	
Chenhe Dong, Guangrun Wang, Hang Xu, Jiefeng Peng, Xiaozhe Ren and Xiaodan Liang ...	1424
<i>Uni-FedRec: A Unified Privacy-Preserving News Recommendation Framework for Model Training and Online Serving</i>	
Tao Qi, Fangzhao Wu, Chuhan Wu, Yongfeng Huang and Xing Xie	1438
<i>Mapping Language to Programs using Multiple Reward Components with Inverse Reinforcement Learning</i>	
Sayan Ghosh and Shashank Srivastava	1449
<i>Topic-Guided Abstractive Multi-Document Summarization</i>	
Peng Cui and Le Hu	1463
<i>An Edge-Enhanced Hierarchical Graph-to-Tree Network for Math Word Problem Solving</i>	
Qinzhao Wu, Qi Zhang and Zhongyu Wei	1473
<i>SciXGen: A Scientific Paper Dataset for Context-Aware Text Generation</i>	
Hong Chen, Hiroya Takamura and Hideki Nakayama	1483
<i>Don't Miss the Potential Customers! Retrieving Similar Ads to Improve User Targeting</i>	
Yi Feng, Ting Wang, Chuanyi Li, Vincent Ng, Jidong Ge, Bin Luo, Yucheng Hu and Xiaopeng Zhang	1493
<i>Cross-lingual Transfer for Text Classification with Dictionary-based Heterogeneous Graph</i>	
Nuttapong Chairatanakul, Noppayut Sriwatanasakdi, Nontawat Charoenphakdee, Xin Liu and Tsuyoshi Murata	1504
<i>Improving Distantly-Supervised Named Entity Recognition with Self-Collaborative Denoising Learning</i>	
Xinghua Zhang, Bowen Yu, Tingwen Liu, Zhenyu Zhang, Jiawei Sheng, Xue Mengge and Hongbo Xu	1518
<i>Entity-Based Semantic Adequacy for Data-to-Text Generation</i>	
Juliette Faille, Albert Gatt and Claire Gardent	1530

<i>MiRANews: Dataset and Benchmarks for Multi-Resource-Assisted News Summarization</i>	
Xinnuo Xu, Ondřej Dušek, Shashi Narayan, Verena Rieser and Ioannis Konstas	1541
<i>A Conditional Generative Matching Model for Multi-lingual Reply Suggestion</i>	
Budhaditya Deb, Guoqing Zheng, Milad Shokouhi and Ahmed Hassan Awadallah	1553
<i>Rethinking Sentiment Style Transfer</i>	
Ping Yu, Yang Zhao, Chunyuan Li and Changyou Chen	1569
<i>HypoGen: Hyperbole Generation with Commonsense and Counterfactual Knowledge</i>	
Yufei Tian, Arvind krishna Sridhar and Nanyun Peng	1583
<i>Profiling News Discourse Structure Using Explicit Subtopic Structures Guided Critics</i>	
Prafulla Kumar Choubey and Ruihong Huang	1594
<i>ProtoInfoMax: Prototypical Networks with Mutual Information Maximization for Out-of-Domain Detection</i>	
Iftitahu Nimah, Meng Fang, Vlado Menkovski and Mykola Pechenizkiy	1606
<i>Learning from Language Description: Low-shot Named Entity Recognition via Decomposed Framework</i>	
Yaqing Wang, Haoda Chu, Chao Zhang and Jing Gao	1618
<i>BERT might be Overkill: A Tiny but Effective Biomedical Entity Linker based on Residual Convolutional Neural Networks</i>	
Tuan Lai, Heng Ji and ChengXiang Zhai	1631
<i>Char2Subword: Extending the Subword Embedding Space Using Robust Character Compositionality</i>	
Gustavo Aguilar, Bryan McCann, Tong Niu, Nazneen Rajani, Nitish Shirish Keskar and Tamar Solorio	1640
<i>Exploring Multitask Learning for Low-Resource Abstractive Summarization</i>	
Ahmed Magooda, Diane Litman and Mohamed Elaraby	1652
<i>Conical Classification For Efficient One-Class Topic Determination</i>	
Sameer Khanna	1662
<i>Improving Dialogue State Tracking with Turn-based Loss Function and Sequential Data Augmentation</i>	
Jarana Manotumruksa, Jeff Dalton, Edgar Meij and Emine Yilmaz	1674
<i>TIAGE: A Benchmark for Topic-Shift Aware Dialog Modeling</i>	
Huiyuan Xie, Zhenghao Liu, Chenyan Xiong, Zhiyuan Liu and Ann Copestake	1684
<i>Optimal Neural Program Synthesis from Multimodal Specifications</i>	
Xi Ye, Qiaochu Chen, Isil Dillig and Greg Durrett	1691
<i>Sent2Span: Span Detection for PICO Extraction in the Biomedical Text without Span Annotations</i>	
Shifeng Liu, Yifang Sun, Bing Li, Wei Wang, Florence T. Bourgeois and Adam G. Dunn	1705
<i>When in Doubt: Improving Classification Performance with Alternating Normalization</i>	
Menglin Jia, Austin Reiter, Ser-Nam Lim, Yoav Artzi and Claire Cardie	1716
<i>APGN: Adversarial and Parameter Generation Networks for Multi-Source Cross-Domain Dependency Parsing</i>	
Ying Li, Meishan Zhang, Zhenghua Li, Min Zhang, Zhefeng Wang, baoxing Huai and Nicholas Jing Yuan	1724

<i>"Let Your Characters Tell Their Story": A Dataset for Character-Centric Narrative Understanding</i> Faeze Brahmaan, Meng Huang, Oyvind Tafjord, Chao Zhao, Mrinmaya Sachan and Snigdha Chaturvedi	1734
<i>Towards Developing a Multilingual and Code-Mixed Visual Question Answering System by Knowledge Distillation</i> Humair Raj Khan, Deepak Gupta and Asif Ekbal	1753
<i>An Iterative Multi-Knowledge Transfer Network for Aspect-Based Sentiment Analysis</i> Yunlong Liang, Fandong Meng, Jinchao Zhang, Yufeng Chen, Jinan Xu and Jie Zhou	1768
<i>Semantic Alignment with Calibrated Similarity for Multilingual Sentence Embedding</i> Jiyeon Ham and Eun-Sol Kim	1781
<i>fBERT: A Neural Transformer for Identifying Offensive Content</i> Diptanu Sarkar, Marcos Zampieri, Tharindu Ranasinghe and Alexander Ororbia	1792
<i>WIKIBIAS: Detecting Multi-Span Subjective Biases in Language</i> Yang Zhong, Jingfeng Yang, Wei Xu and Diyi Yang	1799
<i>UnCIE: Explicitly Leveraging Semantic Similarity to Reduce the Parameters of Word Embeddings</i> Zhi Li, Yuchen Zhai, Chengyu Wang, Minghui Qiu, Kailiang Li and Yin Zhang	1815
<i>Grounded Graph Decoding improves Compositional Generalization in Question Answering</i> Yu Gai, Paras Jain, Wendi Zhang, Joseph Gonzalez, Dawn Song and Ion Stoica	1829
<i>Enhancing Visual Dialog Questioner with Entity-based Strategy Learning and Augmented Guesser</i> Duo Zheng, Zipeng Xu, Fandong Meng, Xiaojie Wang, Jiaan Wang and Jie Zhou	1839
<i>A Pretraining Numerical Reasoning Model for Ordinal Constrained Question Answering on Knowledge Base</i> Yu Feng, Jing Zhang, Gaole He, Wayne Xin Zhao, Lemao Liu, Quan Liu, Cuiping Li and Hong Chen	1852
<i>RoR: Read-over-Read for Long Document Machine Reading Comprehension</i> Jing Zhao, Junwei Bao, Yifan Wang, Yongwei Zhou, Youzheng Wu, Xiaodong He and Bowen Zhou	1862
<i>Span Pointer Networks for Non-Autoregressive Task-Oriented Semantic Parsing</i> Akshat Shrivastava, Pierce Chuang, Arun Babu, Shrey Desai, Abhinav Arora, Alexander Zotov and Ahmed Aly	1873
<i>Language Resource Efficient Learning for Captioning</i> Jia Chen, Yike Wu, Shiwan Zhao and Qin Jin	1887
<i>Translation as Cross-Domain Knowledge: Attention Augmentation for Unsupervised Cross-Domain Segmentation and Labeling Tasks</i> Ruixuan Luo, Yi Zhang, Sishuo Chen and Xu Sun	1896
<i>ContractNLI: A Dataset for Document-level Natural Language Inference for Contracts</i> Yuta Koreeda and Christopher Manning	1907
<i>Japanese Zero Anaphora Resolution Can Benefit from Parallel Texts Through Neural Transfer Learning</i> Masato Umakoshi, Yugo Murawaki and Sadao Kurohashi	1920

<i>Grouped-Attention for Content-Selection and Content-Plan Generation</i>	
Bayu Distiawan Trisedya, Xiaojie Wang, Jianzhong Qi, Rui Zhang and Qingjun Cui	1935
<i>An Explicit-Joint and Supervised-Contrastive Learning Framework for Few-Shot Intent Classification and Slot Filling</i>	
Han Liu, Feng Zhang, Xiaotong Zhang, Siyang Zhao and Xianchao Zhang	1945
<i>Retrieve, Discriminate and Rewrite: A Simple and Effective Framework for Obtaining Affective Response in Retrieval-Based Chatbots</i>	
Xin Lu, Yijian Tian, Yanyan Zhao and Bing Qin	1956
<i>Span Fine-tuning for Pre-trained Language Models</i>	
Rongzhou Bao, Zhuosheng Zhang and Hai Zhao	1970
<i>DIRECT: Direct and Indirect Responses in Conversational Text Corpus</i>	
Junya Takayama, Tomoyuki Kajiwara and Yuki Arase	1980
<i>Retrieval, Analogy, and Composition: A framework for Compositional Generalization in Image Captioning</i>	
Zhan Shi, Hui Liu, Martin Renqiang Min, Christopher Malon, Li Erran Li and Xiaodan Zhu .	1990
<i>TURINGBENCH: A Benchmark Environment for Turing Test in the Age of Neural Text Generation</i>	
Adaku Uchendu, Zeyu Ma, Thai Le, Rui Zhang and Dongwon Lee	2001
<i>Say ‘YES’ to Positivity: Detecting Toxic Language in Workplace Communications</i>	
Meghana Moorthy Bhat, saghar Hosseini, Ahmed Hassan Awadallah, Paul Bennett and Weisheng Li	2017
<i>Natural SQL: Making SQL Easier to Infer from Natural Language Specifications</i>	
Yujian Gan, Xinyun Chen, Jinxia Xie, Matthew Purver, John R. Woodward, John Drake and Qiaofu Zhang	2030
<i>Mitigating Data Scarceness through Data Synthesis, Augmentation and Curriculum for Abstractive Summarization</i>	
Ahmed Magooda and Diane Litman	2043
<i>Self- and Pseudo-self-supervised Prediction of Speaker and Key-utterance for Multi-party Dialogue Reading Comprehension</i>	
Yiyang Li and Hai Zhao	2053
<i>Few-Shot Novel Concept Learning for Semantic Parsing</i>	
Soham Dan, Osbert Bastani and Dan Roth	2064
<i>Compositional Data and Task Augmentation for Instruction Following</i>	
Soham Dan, Xinran Han and Dan Roth	2076
<i>Are Factuality Checkers Reliable? Adversarial Meta-evaluation of Factuality in Summarization</i>	
Yiran Chen, Pengfei Liu and Xipeng Qiu	2082
<i>On the Effects of Transformer Size on In- and Out-of-Domain Calibration</i>	
Soham Dan and Dan Roth	2096
<i>Detecting Polarized Topics Using Partisanship-aware Contextualized Topic Embeddings</i>	
Zihao He, Negar Mokhberian, António Câmara, Andres Abeliuk and Kristina Lerman	2102

<i>GenerativeRE: Incorporating a Novel Copy Mechanism and Pretrained Model for Joint Entity and Relation Extraction</i>	
Jiarun Cao and Sophia Ananiadou	2119
<i>Re-entry Prediction for Online Conversations via Self-Supervised Learning</i>	
Lingzhi Wang, Xingshan Zeng, Huang Hu, Kam-Fai Wong and Daxin Jiang	2127
<i>proScript: Partially Ordered Scripts Generation</i>	
Keisuke Sakaguchi, Chandra Bhagavatula, Ronan Le Bras, Niket Tandon, Peter Clark and Yejin Choi	2138
<i>Speaker Turn Modeling for Dialogue Act Classification</i>	
Zihao He, Leili Tavabi, Kristina Lerman and Mohammad Soleymani	2150
<i>Unsupervised Domain Adaptation Method with Semantic-Structural Alignment for Dependency Parsing</i>	
Boda Lin, Mingzheng Li, Si Li and Yong Luo	2158
<i>Devil’s Advocate: Novel Boosting Ensemble Method from Psychological Findings for Text Classification</i>	
Hwiyeol Jo, Jaeseo Lim and Byoung-Tak Zhang	2168
<i>SideControl: Controlled Open-domain Dialogue Generation via Additive Side Networks</i>	
Wanyu Du and Yangfeng Ji	2175
<i>Is BERT a Cross-Disciplinary Knowledge Learner? A Surprising Finding of Pre-trained Models’ Transferability</i>	
Wei-Tsung Kao and Hung-yi Lee	2195
<i>Geo-BERT Pre-training Model for Query Rewriting in POI Search</i>	
Xiao Liu, Juan Hu, Qi Shen and Huan Chen	2209
<i>Leveraging Bidding Graphs for Advertiser-Aware Relevance Modeling in Sponsored Search</i>	
Shuxian Bi, Chaozhuo Li, Xiao Han, Zheng Liu, Xing Xie, Haizhen Huang and Zengxuan Wen	2215
<i>GPT3Mix: Leveraging Large-scale Language Models for Text Augmentation</i>	
Kang Min Yoo, Dongju Park, Jaewook Kang, Sang-Woo Lee and Woomyoung Park	2225
<i>Context-aware Entity Typing in Knowledge Graphs</i>	
Weiran Pan, Wei Wei and Xian-Ling Mao	2240
<i>Attribute Alignment: Controlling Text Generation from Pre-trained Language Models</i>	
Dian Yu, Zhou Yu and Kenji Sagae	2251
<i>Generate & Rank: A Multi-task Framework for Math Word Problems</i>	
Jianhao Shen, Yichun Yin, Lin Li, Lifeng Shang, Xin Jiang, Ming Zhang and Qun Liu	2269
<i>MIRTT: Learning Multimodal Interaction Representations from Trilinear Transformers for Visual Question Answering</i>	
Junjie Wang, Yatai Ji, Jiaqi Sun, Yujiu Yang and Tetsuya Sakai	2280
<i>UniteD-SRL: A Unified Dataset for Span- and Dependency-Based Multilingual and Cross-Lingual Semantic Role Labeling</i>	
Rocco Tripodi, Simone Conia and Roberto Navigli	2293

<i>Enhancing Dual-Encoders with Question and Answer Cross-Embeddings for Answer Retrieval</i> Yanmeng Wang, Jun Bai, Ye Wang, Jianfei Zhang, Wenge Rong, Zongcheng Ji, Shaojun Wang and Jing Xiao	2306
<i>A Neural Graph-based Local Coherence Model</i> Mohsen Mesgar, Leonardo F. R. Ribeiro and Iryna Gurevych	2316
<i>GiBERT: Enhancing BERT with Linguistic Information using a Lightweight Gated Injection Method</i> Nicole Peinelt, Marek Rei and Maria Liakata	2322
<i>RollingLDA: An Update Algorithm of Latent Dirichlet Allocation to Construct Consistent Time Series from Textual Data</i> Jonas Rieger, Carsten Jentsch and Jörg Rahnenführer	2337
<i>What If Sentence-hood is Hard to Define: A Case Study in Chinese Reading Comprehension</i> Jiawei Wang, Hai Zhao, Yinggong Zhao and Libin Shen	2348
<i>Refining BERT Embeddings for Document Hashing via Mutual Information Maximization</i> Zijing Ou, Qinliang Su, Jianxing Yu, Ruihui Zhao, Yefeng Zheng and Bang Liu	2360
<i>REBEL: Relation Extraction By End-to-end Language generation</i> Pere-Lluís Huguet Cabot and Roberto Navigli	2370
<i>Wine is not v i n. On the Compatibility of Tokenizations across Languages</i> Antonis Maronikolakis, Philipp Dufter and Hinrich Schütze	2382
<i>Temporal Adaptation of BERT and Performance on Downstream Document Classification: Insights from Social Media</i> Paul Röttger and Janet Pierrehumbert	2400
<i>Skim-Attention: Learning to Focus via Document Layout</i> Laura Nguyen, Thomas Scialom, Jacopo Staiano and Benjamin Piwowarski	2413
<i>Attention-based Contrastive Learning for Winograd Schemas</i> Tassilo Klein and Moin Nabi	2428
<i>Give the Truth: Incorporate Semantic Slot into Abstractive Dialogue Summarization</i> Lulu Zhao, Weihao Zeng, Weiran Xu and Jun Guo	2435
<i>Challenges in Detoxifying Language Models</i> Johannes Welbl, Amelia Glaese, Jonathan Uesato, Sumanth Dathathri, John Mellor, Lisa Anne Hendricks, Kirsty Anderson, Pushmeet Kohli, Ben Coppin and Po-Sen Huang	2447
<i>Collecting a Large-Scale Gender Bias Dataset for Coreference Resolution and Machine Translation</i> Shahar Levy, Koren Lazar and Gabriel Stanovsky	2470
<i>Competence-based Curriculum Learning for Multilingual Machine Translation</i> Mingliang Zhang, Fandong Meng, Yunhai Tong and Jie Zhou	2481
<i>Informed Sampling for Diversity in Concept-to-Text NLG</i> Giulio Zhou and Gerasimos Lampouras	2494
<i>Novel Natural Language Summarization of Program Code via Leveraging Multiple Input Representations</i> Fuxiang Chen, Mijung Kim and Jaegul Choo	2510

<i>WikiNEuRal: Combined Neural and Knowledge-based Silver Data Creation for Multilingual NER</i> Simone Tedeschi, Valentino Maiorca, Niccolò Campolungo, Francesco Cecconi and Roberto Navigli	2521
<i>Beyond Grammatical Error Correction: Improving L1-influenced research writing in English using pre-trained encoder-decoder models</i> Gustavo Zomer and Ana Frankenberg-Garcia.....	2534
<i>Classification and Geotemporal Analysis of Quality-of-Life Issues in Tenant Reviews</i> Adam Haber and Zeev Waks.....	2541
<i>Probing Pre-trained Language Models for Semantic Attributes and their Values</i> Meriem Beloucif and Chris Biemann	2554
<i>Uncovering the Limits of Text-based Emotion Detection</i> Nurudin Alvarez-Gonzalez, Andreas Kaltenbrunner and Vicenç Gómez.....	2560
<i>Named Entity Recognition for Entity Linking: What Works and What's Next</i> Simone Tedeschi, Simone Conia, Francesco Cecconi and Roberto Navigli.....	2584
<i>Learning Numeracy: A Simple Yet Effective Number Embedding Approach Using Knowledge Graph</i> Hanyu Duan, Yi Yang and Kar Yan Tam	2597
<i>Weakly Supervised Semantic Parsing by Learning from Mistakes</i> Jiaqi Guo, Jian-Guang Lou, Ting Liu and Dongmei Zhang	2603
<i>CodeQA: A Question Answering Dataset for Source Code Comprehension</i> Chenxiao Liu and Xiaojun Wan	2618
<i>Subword Mapping and Anchoring across Languages</i> Giorgos Vernikos and Andrei Popescu-Belis.....	2633
<i>CDLM: Cross-Document Language Modeling</i> Avi Caciularu, Arman Cohan, Iz Beltagy, Matthew Peters, Arie Cattan and Ido Dagan	2648
<i>Patterns of Polysemy and Homonymy in Contextualised Language Models</i> Janosch Haber and Massimo Poesio	2663
<i>Cross-Lingual Leveled Reading Based on Language-Invariant Features</i> Simin Rao, Hua Zheng and Sujian Li	2677
<i>Controlled Neural Sentence-Level Reframing of News Articles</i> Wei-Fan Chen, Khalid Al Khatib, Benno Stein and Henning Wachsmuth.....	2683
<i>DialogueTRM: Exploring Multi-Modal Emotional Dynamics in a Conversation</i> Yuzhao Mao, Guang Liu, Xiaojie Wang, Weiguo Gao and Xuan Li	2694
<i>Adversarial Examples for Evaluating Math Word Problem Solvers</i> Vivek Kumar, Rishabh Maheshwary and Vikram Pudi	2705
<i>Improving Numerical Reasoning Skills in the Modular Approach for Complex Question Answering on Text</i> Xiao-Yu Guo, Yuan-Fang Li and Gholamreza Haffari	2713
<i>Retrieval Augmented Code Generation and Summarization</i> Md Rizwan Parvez, Wasi Ahmad, Saikat Chakraborty, Baishakhi Ray and Kai-Wei Chang ...	2719

<i>Multilingual Translation via Grafting Pre-trained Language Models</i> Zwei Sun, Mingxuan Wang and Lei Li	2735
<i>AEDA: An Easier Data Augmentation Technique for Text Classification</i> Akbar Karimi, Leonardo Rossi and Andrea Prati	2748
<i>A Comprehensive Comparison of Word Embeddings in Event & Entity Coreference Resolution.</i> Judicael Poumay and Ashwin Ittoo	2755
<i>Wav-BERT: Cooperative Acoustic and Linguistic Representation Learning for Low-Resource Speech Recognition</i> Guolin Zheng, Yubei Xiao, Ke Gong, Pan Zhou, Xiaodan Liang and Liang Lin	2765
<i>Multilingual AMR Parsing with Noisy Knowledge Distillation</i> Deng Cai, Xin Li, Jackie Chun-Sing Ho, Lidong Bing and Wai Lam	2778
<i>Open-Domain Contextual Link Prediction and its Complementarity with Entailment Graphs</i> Mohammad Javad Hosseini, Shay B. Cohen, Mark Johnson and Mark Steedman	2790
<i>Analysis of Language Change in Collaborative Instruction Following</i> Anna Effenberger, Rhia Singh, Eva Yan, Alane Suhr and Yoav Artzi	2803
<i>Counter-Interference Adapter for Multilingual Machine Translation</i> Yaoming Zhu, Jiangtao Feng, Chengqi Zhao, Mingxuan Wang and Lei Li	2812
<i>Progressive Transformer-Based Generation of Radiology Reports</i> Farhad Nooralahzadeh, Nicolas Perez Gonzalez, Thomas Frauenfelder, Koji Fujimoto and Michael Krauthammer	2824
<i>"Be nice to your wife! The restaurants are closed": Can Gender Stereotype Detection Improve Sexism Classification?</i> Patricia Chiril, Farah Benamara and Véronique Moriceau	2833
<i>Automatic Discrimination between Inherited and Borrowed Latin Words in Romance Languages</i> Alina Maria Cristea, Liviu P. Dinu, Simona Georgescu, Mihnea-Lucian Mihai and Ana Sabina Uban	2845
<i>Adapting Language Models for Zero-shot Learning by Meta-tuning on Dataset and Prompt Collections</i> Ruiqi Zhong, Kristy Lee, Zheng Zhang and Dan Klein	2856
<i>Knowledge-Interactive Network with Sentiment Polarity Intensity-Aware Multi-Task Learning for Emotion Recognition in Conversations</i> Yunhe Xie, Kailai Yang, Chengjie Sun, Bingquan Liu and Zhenzhou Ji	2879
<i>Minimizing Annotation Effort via Max-Volume Spectral Sampling</i> Ariadna Quattoni and Xavier Carreras	2890
<i>On the Complementarity between Pre-Training and Back-Translation for Neural Machine Translation</i> Xuebo Liu, Longyue Wang, Derek F. Wong, Liang Ding, Lidia S. Chao, Shuming Shi and Zhaopeng Tu	2900
<i>Lexicon-Based Graph Convolutional Network for Chinese Word Segmentation</i> Kaiyu Huang, Hao Yu, Junpeng Liu, Wei Liu, Jingxiang Cao and Degen Huang	2908
<i>KFCNet: Knowledge Filtering and Contrastive Learning for Generative Commonsense Reasoning</i> Haonan Li, Yeyun Gong, Jian Jiao, Ruofei Zhang, Timothy Baldwin and Nan Duan	2918

<i>Monolingual and Cross-Lingual Acceptability Judgments with the Italian CoLA corpus</i>	
Daniela Trotta, Raffaele Guarasci, Elisa Leonardelli and Sara Tonelli	2929
<i>Hyperbolic Hierarchy-Aware Knowledge Graph Embedding for Link Prediction</i>	
Zhe Pan and Peng Wang	2941
<i>A Discourse-Aware Graph Neural Network for Emotion Recognition in Multi-Party Conversation</i>	
Yang Sun, Nan Yu and Guohong Fu	2949
<i>MeLT: Message-Level Transformer with Masked Document Representations as Pre-Training for Stance Detection</i>	
Matthew Matero, Nikita Soni, Niranjan Balasubramanian and H. Andrew Schwartz	2959
<i>LMSOC: An Approach for Socially Sensitive Pretraining</i>	
Vivek Kulkarni, Shubhanshu Mishra and Aria Haghighi	2967
<i>Extract, Integrate, Compete: Towards Verification Style Reading Comprehension</i>	
Chen Zhang, Yuxuan Lai, Yansong Feng and Dongyan Zhao	2976
<i>Comparing learnability of two dependency schemes: ‘semantic’ (UD) and ‘syntactic’ (SUD)</i>	
Ryszard Tuora, Adam Przepiórkowski and Aleksander Leczkowski	2987
<i>Argumentation-Driven Evidence Association in Criminal Cases</i>	
Yefei Teng and WenHan Chao	2997
<i>Eliminating Sentiment Bias for Aspect-Level Sentiment Classification with Unsupervised Opinion Extraction</i>	
Bo Wang, Tao Shen, Guodong Long, Tianyi Zhou and Yi Chang	3002
<i>Data Efficient Masked Language Modeling for Vision and Language</i>	
Yonatan Bitton, Michael Elhadad, Gabriel Stanovsky and Roy Schwartz	3013
<i>Improving Multilingual Neural Machine Translation with Auxiliary Source Languages</i>	
Weijia Xu, Yuwei Yin, Shuming Ma, Dongdong Zhang and Haoyang Huang	3029
<i>How Does Fine-tuning Affect the Geometry of Embedding Space: A Case Study on Isotropy</i>	
Sara Rajaei and Mohammad Taher Pilehvar	3042
<i>Locality Preserving Sentence Encoding</i>	
Changrong Min, Yonghe Chu, Liang Yang, Bo Xu and Hongfei Lin	3050
<i>Knowledge Representation Learning with Contrastive Completion Coding</i>	
Bo Ouyang, Wenbing Huang, Runfa Chen, Zhixing Tan, Yang Liu, Maosong Sun and Jihong Zhu	3061
<i>Knowledge-Enhanced Evidence Retrieval for Counterargument Generation</i>	
Yohan Jo, Haneul Yoo, JinYeong Bak, Alice Oh, Chris Reed and Eduard Hovy	3074
<i>Investigating Numeracy Learning Ability of a Text-to-Text Transfer Model</i>	
Kuntal Kumar Pal and Chitta Baral	3095
<i>Modeling Mathematical Notation Semantics in Academic Papers</i>	
Hwiyeol Jo, Dongyeop Kang, Andrew Head and Marti A. Hearst	3102

<i>Unpacking the Interdependent Systems of Discrimination: Ableist Bias in NLP Systems through an Intersectional Lens</i>	
Saad Hassan, Matt Huenerfauth and Cecilia Ovesdotter Alm	3116
<i>Constructing Emotional Consensus and Utilizing Unpaired Data for Empathetic Dialogue Generation</i>	
Lei Shen, Jinchao Zhang, Jiao Ou, Xiaofang Zhao and Jie Zhou	3124
<i>Automatic rule generation for time expression normalization</i>	
Wentao Ding, Jianhao Chen, Jinmao Li and Yuzhong Qu	3135
<i>RW-KD: Sample-wise Loss Terms Re-Weighting for Knowledge Distillation</i>	
Peng Lu, Abbas Ghaddar, Ahmad Rashid, Mehdi Rezagholizadeh, Ali Ghodsi and Philippe Langlais	3145
<i>Visual Cues and Error Correction for Translation Robustness</i>	
Zhenhao Li, Marek Rei and Lucia Specia	3153
<i>Beyond the Tip of the Iceberg: Assessing Coherence of Text Classifiers</i>	
Shane Storks and Joyce Chai	3169
<i>Does Pretraining for Summarization Require Knowledge Transfer?</i>	
Kundan Krishna, Jeffrey Bigham and Zachary C. Lipton	3178
<i>Bandits Don't Follow Rules: Balancing Multi-Facet Machine Translation with Multi-Armed Bandits</i>	
Julia Kreutzer, David Vilar and Artem Sokolov	3190
<i>Sometimes We Want Ungrammatical Translations</i>	
Prasanna Parthasarathi, Koustuv Sinha, Joelle Pineau and Adina Williams	3205
<i>An animated picture says at least a thousand words: Selecting Gif-based Replies in Multimodal Dialog</i>	
Xingyao Wang and David Jurgens	3228
<i>SciCap: Generating Captions for Scientific Figures</i>	
Ting-Yao Hsu, C Lee Giles and Ting-Hao Huang	3258
<i>SentNoB: A Dataset for Analysing Sentiment on Noisy Bangla Texts</i>	
Khondoker Ittehadul Islam, Sudipta Kar, Md Saiful Islam and Mohammad Ruhul Amin	3265
<i>Translate & Fill: Improving Zero-Shot Multilingual Semantic Parsing with Synthetic Data</i>	
Massimo Nicosia, Zhongdi Qu and Yasemin Altun	3272
<i>NewsBERT: Distilling Pre-trained Language Model for Intelligent News Application</i>	
Chuhan Wu, Fangzhao Wu, Yang Yu, Tao Qi, Yongfeng Huang and Qi Liu	3285
<i>SD-QA: Spoken Dialectal Question Answering for the Real World</i>	
Fahim Faisal, Sharlina Keshava, Md Mahfuz Ibn Alam and Antonios Anastasopoulos	3296
<i>The Low-Resource Double Bind: An Empirical Study of Pruning for Low-Resource Machine Translation</i>	
Orevaoghene Ahia, Julia Kreutzer and Sara Hooker	3316
<i>Transformer over Pre-trained Transformer for Neural Text Segmentation with Enhanced Topic Coherence</i>	
Kelvin Lo, Yuan Jin, Weicong Tan, Ming Liu, Lan Du and Wray Buntine	3334
<i>Self-Supervised Neural Topic Modeling</i>	
Seyed Ali Bahrainian, Martin Jaggi and Carsten Eickhoff	3341

<i>Coreference-aware Surprisal Predicts Brain Response</i>	
Evan Jaffe, Byung-Doh Oh and William Schuler	3351
<i>Distilling the Knowledge of Large-scale Generative Models into Retrieval Models for Efficient Open-domain Conversation</i>	
Beomsu Kim, Seokjun Seo, Seungju Han, Enkhbayar Erdenee and Buru Chang	3357
<i>Modeling Users and Online Communities for Abuse Detection: A Position on Ethics and Explainability</i>	
Pushkar Mishra, Helen Yannakoudakis and Ekaterina Shutova	3374
<i>Detecting Community Sensitive Norm Violations in Online Conversations</i>	
Chan Young Park, Julia Mendelsohn, Karthik Radhakrishnan, Kinjal Jain, Tushar Kanakagiri, David Jurgens and Yulia Tsvetkov	3386
<i>SupCL-Seq: Supervised Contrastive Learning for Downstream Optimized Sequence Representations</i>	
Hooman Sedghamiz, Shivam Raval, Enrico Santus, Tuka Alhanai and Mohammad Ghassemi	3398
<i>mDAPT: Multilingual Domain Adaptive Pretraining in a Single Model</i>	
Rasmus Kær Jørgensen, Mareike Hartmann, Xiang Dai and Desmond Elliott	3404
<i>COSMic: A Coherence-Aware Generation Metric for Image Descriptions</i>	
Mert Inan, Piyush Sharma, Baber Khalid, Radu Soricut, Matthew Stone and Malihe Alikhani	3419
<i>Relation-Guided Pre-Training for Open-Domain Question Answering</i>	
Ziniu Hu, Yizhou Sun and Kai-Wei Chang	3431
<i>MURAL: Multimodal, Multitask Representations Across Languages</i>	
Aashi Jain, Mandy Guo, Krishna Srinivasan, Ting Chen, Sneha Kudugunta, Chao Jia, Yinfei Yang and Jason Baldridge	3449
<i>AStitchInLanguageModels: Dataset and Methods for the Exploration of Idiomaticity in Pre-Trained Language Models</i>	
Harish Tayyar Madabushi, Edward Gow-Smith, Carolina Scarton and Aline Villavicencio	3464
<i>Refine and Imitate: Reducing Repetition and Inconsistency in Persuasion Dialogues via Reinforcement Learning and Human Demonstration</i>	
Weiyang Shi, Yu Li, Saurav Sahay and Zhou Yu	3478
<i>A Computational Exploration of Pejorative Language in Social Media</i>	
Liviu P. Dinu, Ioan-Bogdan Iordache, Ana Sabina Uban and Marcos Zampieri	3493
<i>Evidence-based Fact-Checking of Health-related Claims</i>	
Mourad Sarrouti, Asma Ben Abacha, Yassine Mrabet and Dina Demner-Fushman	3499
<i>Learning and Analyzing Generation Order for Undirected Sequence Models</i>	
Yichen Jiang and Mohit Bansal	3513
<i>Automatic Bilingual Markup Transfer</i>	
Thomas Zenkel, Joern Wuebker and John DeNero	3524
<i>Exploring a Unified Sequence-To-Sequence Transformer for Medical Product Safety Monitoring in Social Media</i>	
Shivam Raval, Hooman Sedghamiz, Enrico Santus, Tuka Alhanai, Mohammad Ghassemi and Emanuele Chersoni	3534

<i>Disentangling Generative Factors in Natural Language with Discrete Variational Autoencoders</i> Giangiacomo Mercatali and André Freitas	3547
<i>MSD: Saliency-aware Knowledge Distillation for Multimodal Understanding</i> Woojeong Jin, Maziar Sanjabi, Shaoliang Nie, Liang Tan, Xiang Ren and Hamed Firooz	3557
<i>Do UD Trees Match Mention Spans in Coreference Annotations?</i> Martin Popel, Zdeněk Žabokrtský, Anna Nedoluzhko, Michal Novák and Daniel Zeman	3570
<i>Beyond Distillation: Task-level Mixture-of-Experts for Efficient Inference</i> Sneha Kudugunta, Yanping Huang, Ankur Bapna, Maxim Krikun, Dmitry Lepikhin, Minh-Thang Luong and Orhan Firat	3577
<i>TAG: Gradient Attack on Transformer-based Language Models</i> Jieren Deng, Yijue Wang, Ji Li, Chenghong Wang, Chao Shang, Hang Liu, Sanguthevar Rajasekaran and Caiwen Ding	3600
<i>Generating Realistic Natural Language Counterfactuals</i> Marcel Robeer, Floris Bex and Ad Feelders	3611
<i>Unsupervised Chunking as Syntactic Structure Induction with a Knowledge-Transfer Approach</i> Anup Anand Deshmukh, Qianqiu Zhang, Ming Li, Jimmy Lin and Lili Mou	3626
<i>Model-based analysis of brain activity reveals the hierarchy of language in 305 subjects</i> Charlotte Caucheteux, Alexandre Gramfort and Jean-Remi King	3635
<i>Gated Transformer for Robust De-noised Sequence-to-Sequence Modelling</i> Ayan Sengupta, Amit Kumar, Sourabh Kumar Bhattacharjee and Suman Roy	3645
<i>Token-wise Curriculum Learning for Neural Machine Translation</i> Chen Liang, Haoming Jiang, Xiaodong Liu, Pengcheng He, Weizhu Chen, Jianfeng Gao and Tuo Zhao	3658
<i>RelDiff: Enriching Knowledge Graph Relation Representations for Sensitivity Classification</i> Hitarth Narvala, Graham McDonald and Iadh Ounis	3671
<i>Post-Editing Extractive Summaries by Definiteness Prediction</i> Jad Kabbara and Jackie Chi Kit Cheung	3682
<i>Leveraging Pretrained Models for Automatic Summarization of Doctor-Patient Conversations</i> Longxiang Zhang, Renato Negrinho, Arindam Ghosh, Vasudevan Jagannathan, Hamid Reza Hasanzadeh, Thomas Schaaf and Matthew R. Gormley	3693
<i>Distilling Knowledge for Empathy Detection</i> Mahshid Hosseini and Cornelia Caragea	3713
<i>Adapting Entities across Languages and Cultures</i> Denis Peskov, Viktor Hangya, Jordan Boyd-Graber and Alexander Fraser	3725
<i>ODIST: Open World Classification via Distributionally Shifted Instances</i> Lei Shu, Yassine Benajiba, Saab Mansour and Yi Zhang	3751
<i>LAMAD: A Linguistic Attentional Model for Arabic Text Diacritization</i> Raed Al-Sabri and Jianliang Gao	3757

<i>Sequence-to-Lattice Models for Fast Translation</i> Yuntian Deng and Alexander Rush	3765
<i>Towards Realistic Single-Task Continuous Learning Research for NER</i> Justin Payan, Yuval Merhav, He Xie, Satyapriya Krishna, Anil Ramakrishna, Mukund Sridhar and Rahul Gupta	3773
<i>Retrieval Augmentation Reduces Hallucination in Conversation</i> Kurt Shuster, Spencer Poff, Moya Chen, Douwe Kiela and Jason Weston	3784
<i>Towards Automatic Bias Detection in Knowledge Graphs</i> Daphna Keidar, Mian Zhong, Ce Zhang, Yash Raj Shrestha and Bibek Paudel	3804
<i>Searching for More Efficient Dynamic Programs</i> Tim Vieira, Ryan Cotterell and Jason Eisner	3812
<i>Revisiting Robust Neural Machine Translation: A Transformer Case Study</i> Peyman Passban, Puneeth Saladi and Qun Liu	3831
<i>Can NLI Models Verify QA Systems' Predictions?</i> Jifan Chen, Eunsol Choi and Greg Durrett	3841
<i>Parameter-Efficient Domain Knowledge Integration from Multiple Sources for Biomedical Pre-trained Language Models</i> Qiuhaio Lu, Dejing Dou and Thien Huu Nguyen	3855
<i>Uncovering Implicit Gender Bias in Narratives through Commonsense Inference</i> Tenghao Huang, Faeze Brahman, Vered Shwartz and Snigdha Chaturvedi	3866
<i>Contrastive Document Representation Learning with Graph Attention Networks</i> Peng Xu, Xinchu Chen, Xiaofei Ma, Zhiheng Huang and Bing Xiang	3874
<i>Convex Aggregation for Opinion Summarization</i> Hayate Iso, Xiaolan Wang, Yoshihiko Suhara, Stefanos Angelidis and Wang-Chiew Tan	3885
<i>Using Optimal Transport as Alignment Objective for fine-tuning Multilingual Contextualized Embed- dings</i> Sawsan Alqahtani, Garima Lalwani, Yi Zhang, Salvatore Romeo and Saab Mansour	3904
<i>Uncertainty-Aware Machine Translation Evaluation</i> Taisiya Glushkova, Chrysoula Zerva, Ricardo Rei and André F. T. Martins	3920
<i>Neural Unification for Logic Reasoning over Natural Language</i> Gabriele Picco, Thanh Lam Hoang, Marco Luca Sbordio and Vanessa Lopez	3939
<i>From None to Severe: Predicting Severity in Movie Scripts</i> Yigeng Zhang, Mahsa Shafaei, Fabio Gonzalez and Thamar Solorio	3951
<i>Benchmarking Meta-embeddings: What Works and What Does Not</i> Iker García-Ferrero, Rodrigo Agerri and German Rigau	3957
<i>A Plug-and-Play Method for Controlled Text Generation</i> Damian Pascual, Beni Egressy, Clara Meister, Ryan Cotterell and Roger Wattenhofer	3973
<i>A Corpus-based Syntactic Analysis of Two-termed Unlike Coordination</i> Julie Kallini and Christiane Fellbaum	3998

<i>Weakly Supervised Contrastive Learning for Chest X-Ray Report Generation</i> An Yan, Zexue He, Xing Lu, Jiang Du, Eric Chang, Amilcare Gentili, Julian McAuley and Chun-Nan Hsu	4009
<i>NUANCED: Natural Utterance Annotation for Nuanced Conversation with Estimated Distributions</i> Zhiyu Chen, Honglei Liu, Hu Xu, Seungwhan Moon, Hao Zhou and Bing Liu	4016
<i>Table-based Fact Verification With Saliency-aware Learning</i> Fei Wang, Kexuan Sun, Jay Pujara, Pedro Szekely and Muhao Chen	4025
<i>Detecting Frames in News Headlines and Lead Images in U.S. Gun Violence Coverage</i> Isidora Tourni, Lei Guo, Taufiq Husada Daryanto, Fabian Zhafransyah, Edward Edberg Halim, Mona Jalal, Boqi Chen, Sha Lai, Hengchang Hu, Margrit Betke, Prakash Ishwar and Derry Tanti Wijaya	4037
<i>Multi-task Learning to Enable Location Mention Identification in the Early Hours of a Crisis Event</i> Sarthak Khanal and Doina Caragea	4051
<i>Graph-Based Decoding for Task Oriented Semantic Parsing</i> Jeremy Cole, Nanjiang Jiang, Panupong Pasupat, Luheng He and Peter Shaw	4057
<i>Expected Validation Performance and Estimation of a Random Variable's Maximum</i> Jesse Dodge, Suchin Gururangan, Dallas Card, Roy Schwartz and Noah A. Smith	4066
<i>How May I Help You? Using Neural Text Simplification to Improve Downstream NLP Tasks</i> Hoang Van, Zheng Tang and Mihai Surdeanu	4074
<i>Subformer: Exploring Weight Sharing for Parameter Efficiency in Generative Transformers</i> Machel Reid, Edison Marrese-Taylor and Yutaka Matsuo	4081
<i>Leveraging Information Bottleneck for Scientific Document Summarization</i> Jiaxin Ju, Ming Liu, Huan Yee Koh, Yuan Jin, Lan Du and Shirui Pan	4091
<i>Reconsidering the Past: Optimizing Hidden States in Language Models</i> Davis Yoshida and Kevin Gimpel	4099
<i>Attend, Memorize and Generate: Towards Faithful Table-to-Text Generation in Few Shots</i> Wenting Zhao, Ye Liu, Yao Wan and Philip Yu	4106
<i>ARCH: Efficient Adversarial Regularized Training with Caching</i> Simiao Zuo, Chen Liang, Haoming Jiang, Pengcheng He, Xiaodong Liu, Jianfeng Gao, Weizhu Chen and Tuo Zhao	4118
<i>Probing Commonsense Explanation in Dialogue Response Generation</i> Pei Zhou, Pegah Jandaghi, Hyundong Cho, Bill Yuchen Lin, Jay Pujara and Xiang Ren	4132
<i>NOAHQA: Numerical Reasoning with Interpretable Graph Question Answering Dataset</i> Qiyuan Zhang, Lei Wang, Sicheng Yu, Shuohang Wang, Yang Wang, Jing Jiang and Ee-Peng Lim	4147
<i>Textual Time Travel: A Temporally Informed Approach to Theory of Mind</i> Akshatha Arodi and Jackie Chi Kit Cheung	4162
<i>Detect and Perturb: Neutral Rewriting of Biased and Sensitive Text via Gradient-based Decoding</i> Zexue He, Bodhisattwa Prasad Majumder and Julian McAuley	4173

<i>HyperExpan: Taxonomy Expansion with Hyperbolic Representation Learning</i>	
Mingyu Derek Ma, Muhao Chen, Te-Lin Wu and Nanyun Peng	4182
<i>Want To Reduce Labeling Cost? GPT-3 Can Help</i>	
Shuohang Wang, Yang Liu, Yichong Xu, Chenguang Zhu and Michael Zeng	4195
<i>Written Justifications are Key to Aggregate Crowdsourced Forecasts</i>	
Saketh Kotamraju and Eduardo Blanco	4206
<i>Cleaning Dirty Books: Post-OCR Processing for Previously Scanned Texts</i>	
Allen Kim, Charuta Pethe, Naoya Inoue and Steve Skiena	4217
<i>Bag of Tricks for Optimizing Transformer Efficiency</i>	
Ye Lin, Yanyang Li, Tong Xiao and Jingbo Zhu	4227
<i>Non-Parametric Unsupervised Domain Adaptation for Neural Machine Translation</i>	
Xin Zheng, Zhirui Zhang, Shujian Huang, Boxing Chen, Jun Xie, Weihua Luo and Jiajun Chen	4234
<i>The Topic Confusion Task: A Novel Evaluation Scenario for Authorship Attribution</i>	
Malik Altakrori, Jackie Chi Kit Cheung and Benjamin C. M. Fung	4242
<i>Micromodels for Efficient, Explainable, and Reusable Systems: A Case Study on Mental Health</i>	
Andrew Lee, Jonathan K. Kummerfeld, Larry An and Rada Mihalcea	4257
<i>Discovering Explanatory Sentences in Legal Case Decisions Using Pre-trained Language Models</i>	
Jaromir Savelka and Kevin Ashley	4273
<i>FCM: A Fine-grained Comparison Model for Multi-turn Dialogue Reasoning</i>	
Xu Wang, Hainan Zhang, Shuai Zhao, Yanyan Zou, Hongshen Chen, Zhuoye Ding, Bo Cheng and Yanyan Lan	4284
<i>Reference-based Weak Supervision for Answer Sentence Selection using Web Data</i>	
Vivek Krishnamurthy, Thuy Vu and Alessandro Moschitti	4294
<i>A Deep Decomposable Model for Disentangling Syntax and Semantics in Sentence Representation</i>	
Dingcheng Li, Hongliang Fei, Shaogang Ren and Ping Li	4300
<i>Improved Word Sense Disambiguation with Enhanced Sense Representations</i>	
Yang Song, Xin Cai Ong, Hwee Tou Ng and Qian Lin	4311
<i>Rethinking Zero-shot Neural Machine Translation: From a Perspective of Latent Variables</i>	
Weizhi Wang, Zhirui Zhang, Yichao Du, Boxing Chen, Jun Xie and Weihua Luo	4321
<i>FastCorrect 2: Fast Error Correction on Multiple Candidates for Automatic Speech Recognition</i>	
Yichong Leng, Xu Tan, Rui Wang, Linchen Zhu, Jin Xu, Wenjie Liu, Linqun Liu, Xiang-Yang Li, Tao Qin, Edward Lin and Tie-Yan Liu	4328
<i>Task-Oriented Clustering for Dialogues</i>	
Chenxu Lv, Hengtong Lu, Shuyu Lei, Huixing Jiang, Wei Wu, Caixia Yuan and Xiaojie Wang	4338
<i>Mitigating Data Poisoning in Text Classification with Differential Privacy</i>	
Chang Xu, Jun Wang, Francisco Guzmán, Benjamin Rubinstein and Trevor Cohn	4348
<i>Does Vision-and-Language Pretraining Improve Lexical Grounding?</i>	
Tian Yun, Chen Sun and Ellie Pavlick	4357

<i>Character-based PCFG Induction for Modeling the Syntactic Acquisition of Morphologically Rich Languages</i>	
Lifeng Jin, Byung-Doh Oh and William Schuler	4367
<i>Block-wise Word Embedding Compression Revisited: Better Weighting and Structuring</i>	
Jong-Ryul Lee, Yong-Ju Lee and Yong-Hyuk Moon	4379
<i>Switch Point biased Self-Training: Re-purposing Pretrained Models for Code-Switching</i>	
Parul Chopra, Sai Krishna Rallabandi, Alan W Black and Khyathi Raghavi Chandu	4389
<i>Influence Tuning: Demoting Spurious Correlations via Instance Attribution and Instance-Driven Updates</i>	
Xiaochuang Han and Yulia Tsvetkov	4398
<i>Learning Task Sampling Policy for Multitask Learning</i>	
Dhanasekar Sundararaman, Henry Tsai, Kuang-Huei Lee, Iulia Turc and Lawrence Carin	4410
<i>Competing Independent Modules for Knowledge Integration and Optimization</i>	
Parsa Bagherzadeh and Sabine Bergler	4416
<i>An Exploratory Study on Long Dialogue Summarization: What Works and What’s Next</i>	
Yusen Zhang, Ansong Ni, Tao Yu, Rui Zhang, Chenguang Zhu, Budhaditya Deb, Asli Celikyilmaz, Ahmed Hassan Awadallah and Dragomir Radev	4426
<i>Improving Text Auto-Completion with Next Phrase Prediction</i>	
Dong-Ho Lee, Zhiqiang Hu and Roy Ka-Wei Lee	4434
<i>MOMENTA: A Multimodal Framework for Detecting Harmful Memes and Their Targets</i>	
Shraman Pramanick, Shivam Sharma, Dimitar Dimitrov, Md. Shad Akhtar, Preslav Nakov and Tanmoy Chakraborty	4439
<i>NICE: Neural Image Commenting with Empathy</i>	
Kezhen Chen, Qiuyuan Huang, Daniel McDuff, Xiang Gao, Hamid Palangi, Jianfeng Wang, Kenneth Forbus and Jianfeng Gao	4456
<i>HACovGNN: Hierarchical Attention Based Convolutional Graph Neural Network for Code Documentation Generation in Jupyter Notebooks</i>	
Xuye Liu, Dakuo Wang, April Wang, Yufang Hou and Lingfei Wu	4473
<i>A multilabel approach to morphosyntactic probing</i>	
Naomi Shapiro, Amandalynne Paullada and Shane Steinert-Threlkeld	4486
<i>Co-Teaching Student-Model through Submission Results of Shared Task</i>	
Kouta Nakayama, Shuhei Kurita, Akio Kobayashi, Yukino Baba and Satoshi Sekine	4525
<i>KLMo: Knowledge Graph Enhanced Pretrained Language Model with Fine-Grained Relationships</i>	
Lei He, Suncong Zheng, Tao Yang and Feng Zhang	4536
<i>Do We Know What We Don’t Know? Studying Unanswerable Questions beyond SQuAD 2.0</i>	
Elior Sulem, Jamaal Hay and Dan Roth	4543
<i>Glyph Enhanced Chinese Character Pre-Training for Lexical Sememe Prediction</i>	
Boer Lyu, Lu Chen and Kai Yu	4549
<i>Active Learning for Rumor Identification on Social Media</i>	
Parsa Farinneya, Mohammad Mahdi Abdollah Pour, Sardar Hamidian and Mona Diab	4556

<i>Cross-Domain Data Integration for Named Entity Disambiguation in Biomedical Text</i>	
Maya Varma, Laurel Orr, Sen Wu, Megan Leszczynski, Xiao Ling and Christopher Ré	4566
<i>Self-Training using Rules of Grammar for Few-Shot NLU</i>	
Joonghyuk Hahn, Hyunjoon Cheon, Kyuyeol Han, Cheongjae Lee, Junseok Kim and Yo-Sub Han	4576
<i>Aspect-based Sentiment Analysis in Question Answering Forums</i>	
Wenxuan Zhang, Yang Deng, Xin Li, Lidong Bing and Wai Lam	4582
<i>ForumSum: A Multi-Speaker Conversation Summarization Dataset</i>	
Misha Khalman, Yao Zhao and Mohammad Saleh	4592
<i>Question Answering over Electronic Devices: A New Benchmark Dataset and a Multi-Task Learning based QA Framework</i>	
Abhilash Nandy, Soumya Sharma, Shubham Maddhashiya, Kapil Sachdeva, Pawan Goyal and Niloy Ganguly	4600
<i>Comprehensive Punctuation Restoration for English and Polish</i>	
Michał Pogoda and Tomasz Walkowiak	4610
<i>Syntactically Diverse Adversarial Network for Knowledge-Grounded Conversation Generation</i>	
Fuwei Cui, Hui Di, Hongjie Ren, Kazushige Ouchi, Ze Liu and Jinan Xu	4620
<i>QACE: Asking Questions to Evaluate an Image Caption</i>	
Hwanhee Lee, Thomas Scialom, Seunghyun Yoon, Franck Dernoncourt and Kyomin Jung . . .	4631
<i>Secoco: Self-Correcting Encoding for Neural Machine Translation</i>	
Tao Wang, Chengqi Zhao, Mingxuan Wang, Lei Li, Hang Li and Deyi Xiong	4639
<i>Simple or Complex? Complexity-controllable Question Generation with Soft Templates and Deep Mixture of Experts Model</i>	
Sheng Bi, Xiya Cheng, Yuan-Fang Li, Lizhen Qu, Shirong Shen, Guilin Qi, Lu Pan and Yinlin Jiang	4645
<i>Predicting Anti-Asian Hateful Users on Twitter during COVID-19</i>	
Jisun An, Haewoon Kwak, Claire Seungeun Lee, Bogang Jun and Yong-Yeol Ahn	4655
<i>Fine-grained Typing of Emerging Entities in Microblogs</i>	
Satoshi Akasaki, Naoki Yoshinaga and Masashi Toyoda	4667
<i>Data-Efficient Language Shaped Few-shot Image Classification</i>	
Zhenwen Liang and Xiangliang Zhang	4680
<i>Beyond Glass-Box Features: Uncertainty Quantification Enhanced Quality Estimation for Neural Machine Translation</i>	
Ke Wang, Yangbin Shi, Jiayi Wang, Yuqi Zhang, Yu Zhao and Xiaolin Zheng	4687
<i>Fight Fire with Fire: Fine-tuning Hate Detectors using Large Samples of Generated Hate Speech</i>	
Tomer Wullach, Amir Adler and Einat Minkov	4699
<i>AutoEQA: Auto-Encoding Questions for Extractive Question Answering</i>	
Stalin Varanasi, Saadullah Amin and Guenter Neumann	4706
<i>A Multi-label Multi-hop Relation Detection Model based on Relation-aware Sequence Generation</i>	
Linhai Zhang, Deyu Zhou, Chao Lin and Yulan He	4713

<i>Don't Discard All the Biased Instances: Investigating a Core Assumption in Dataset Bias Mitigation Techniques</i>	
Hossein Amirkhani and Mohammad Taher Pilehvar	4720
<i>Stacked AMR Parsing with Silver Data</i>	
Qingrong Xia, Zhenghua Li, Rui Wang and Min Zhang	4729
<i>Speculative Sampling in Variational Autoencoders for Dialogue Response Generation</i>	
Shoetsu Sato, Naoki Yoshinaga, Masashi Toyoda and Masaru Kitsuregawa	4739
<i>Perceived and Intended Sarcasm Detection with Graph Attention Networks</i>	
Joan Plepi and Lucie Flek	4746
<i>Contrastive Representation Learning for Exemplar-Guided Paraphrase Generation</i>	
Haoran Yang, Wai Lam and Piji Li	4754
<i>MAD-G: Multilingual Adapter Generation for Efficient Cross-Lingual Transfer</i>	
Alan Ansell, Edoardo Maria Ponti, Jonas Pfeiffer, Sebastian Ruder, Goran Glavaš, Ivan Vulić and Anna Korhonen	4762
<i>Sustainable Modular Debiasing of Language Models</i>	
Anne Lauscher, Tobias Lueken and Goran Glavaš	4782
<i>A Divide-And-Conquer Approach for Multi-label Multi-hop Relation Detection in Knowledge Base Question Answering</i>	
Deyu Zhou, Yanzheng Xiang, Linhai Zhang, Chenchen Ye, Qian-Wen Zhang and Yunbo Cao	4798
<i>Counterfactual Adversarial Learning with Representation Interpolation</i>	
Wei Wang, Boxin Wang, Ning Shi, Jinfeng Li, Bingyu Zhu, Xiangyu Liu and Rong Zhang	4809
<i>'Just What do You Think You're Doing, Dave?' A Checklist for Responsible Data Use in NLP</i>	
Anna Rogers, Timothy Baldwin and kobi leins	4821
<i>Counter-Contrastive Learning for Language GANs</i>	
Yekun Chai, Haidong Zhang, Qiyue Yin and Junge Zhang	4834
<i>Incorporating Circumstances into Narrative Event Prediction</i>	
Shichao Wang, Xiangrui Cai, HongBin Wang and Xiaojie Yuan	4840
<i>MultiFix: Learning to Repair Multiple Errors by Optimal Alignment Learning</i>	
HyeonTae Seo, Yo-Sub Han and Sang-Ki Ko	4850
<i>HOTTER: Hierarchical Optimal Topic Transport with Explanatory Context Representations</i>	
Sabine Wehnert, Christian Scheel, Simona Szakács-Behling, Maret Nieländer, Patrick Mielke and Ernesto William De Luca	4856
<i>Grammatical Error Correction with Contrastive Learning in Low Error Density Domains</i>	
Hannan Cao, Wenmian Yang and Hwee Tou Ng	4867
<i>Improving Unsupervised Commonsense Reasoning Using Knowledge-Enabled Natural Language Inference</i>	
Canming Huang, Weinan He and Yongmei Liu	4875
<i>Does Putting a Linguist in the Loop Improve NLU Data Collection?</i>	
Alicia Parrish, William Huang, Omar Agha, Soo-Hwan Lee, Nikita Nangia, Alexia Warstadt, Karmanya Aggarwal, Emily Allaway, Tal Linzen and Samuel R. Bowman	4886

<i>Tiered Reasoning for Intuitive Physics: Toward Verifiable Commonsense Language Understanding</i> Shane Storks, Qiaozi Gao, Yichi Zhang and Joyce Chai	4902
<i>Making Heads and Tails of Models with Marginal Calibration for Sparse Tagsets</i> Michael Kranzlein, Nelson F. Liu and Nathan Schneider.....	4919
<i>GeDi: Generative Discriminator Guided Sequence Generation</i> Ben Krause, Akhilesh Deepak Gotmare, Bryan McCann, Nitish Shirish Keskar, Shafiq Joty, Richard Socher and Nazneen Fatema Rajani	4929