

Conference on Empirical Methods in Natural Language Processing (EMNLP 2021)

Punta Cana, Dominican Republic and Online
7 – 11 November 2021

Volume 1 of 15

ISBN: 978-1-7138-9161-1

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. (2024)

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

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

Phone: 1-570-476-8006

Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

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

Table of Contents

<i>AlignNART: Non-autoregressive Neural Machine Translation by Jointly Learning to Estimate Alignment and Translate</i>	
Jongyoon Song, Sungwon Kim and Sungroh Yoon	1
<i>Zero-Shot Cross-Lingual Transfer of Neural Machine Translation with Multilingual Pretrained Encoders</i>	
Guanhua Chen, Shuming Ma, Yun Chen, Li Dong, Dongdong Zhang, Jia Pan, Wenping Wang and Furu Wei	15
<i>ERNIE-M: Enhanced Multilingual Representation by Aligning Cross-lingual Semantics with Monolingual Corpora</i>	
Xuan Ouyang, Shuohuan Wang, Chao Pang, Yu Sun, Hao Tian, Hua Wu and Haifeng Wang	27
<i>Cross Attention Augmented Transducer Networks for Simultaneous Translation</i>	
Dan Liu, Mengge Du, Xiaoxi Li, Ya Li and Enhong Chen	39
<i>Translating Headers of Tabular Data: A Pilot Study of Schema Translation</i>	
Kunrui Zhu, Yan Gao, Jiaqi Guo and Jian-Guang Lou	56
<i>Towards Making the Most of Dialogue Characteristics for Neural Chat Translation</i>	
Yunlong Liang, Chulun Zhou, Fandong Meng, Jinan Xu, Yufeng Chen, Jinsong Su and Jie Zhou	67
<i>Low-Resource Dialogue Summarization with Domain-Agnostic Multi-Source Pretraining</i>	
Yicheng Zou, Bolin Zhu, Xingwu Hu, Tao Gui and Qi Zhang	80
<i>Controllable Neural Dialogue Summarization with Personal Named Entity Planning</i>	
Zhengyuan Liu and Nancy Chen	92
<i>Fine-grained Factual Consistency Assessment for Abstractive Summarization Models</i>	
Sen Zhang, Jianwei Niu and Chuyuan Wei	107
<i>Decision-Focused Summarization</i>	
Chao-Chun Hsu and Chenhao Tan	117
<i>Multiplex Graph Neural Network for Extractive Text Summarization</i>	
Baoyu Jing, Zeyu You, Tao Yang, Wei Fan and Hanghang Tong	133
<i>A Thorough Evaluation of Task-Specific Pretraining for Summarization</i>	
Sascha Rothe, Joshua Maynez and Shashi Narayan	140
<i>HETFORMER: Heterogeneous Transformer with Sparse Attention for Long-Text Extractive Summarization</i>	
Ye Liu, Jianguo Zhang, Yao Wan, Congying Xia, Lifang He and Philip Yu	146
<i>Unsupervised Keyphrase Extraction by Jointly Modeling Local and Global Context</i>	
Xinnian Liang, Shuangzhi Wu, Mu Li and Zhoujun Li	155
<i>Distantly Supervised Relation Extraction using Multi-Layer Revision Network and Confidence-based Multi-Instance Learning</i>	
Xiangyu Lin, Tianyi Liu, Weijia Jia and Zhiguo Gong	165
<i>Logic-level Evidence Retrieval and Graph-based Verification Network for Table-based Fact Verification</i>	
Qi Shi, Yu Zhang, Qingyu Yin and Ting Liu	175

<i>A Partition Filter Network for Joint Entity and Relation Extraction</i>	
Zhiheng Yan, Chong Zhang, Jinlan Fu, Qi Zhang and Zhongyu Wei	185
<i>TEBNER: Domain Specific Named Entity Recognition with Type Expanded Boundary-aware Network</i>	
Zheng Fang, Yanan Cao, Tai Li, Ruipeng Jia, Fang Fang, Yanmin Shang and Yuhai Lu	198
<i>Beta Distribution Guided Aspect-aware Graph for Aspect Category Sentiment Analysis with Affective Knowledge</i>	
Bin Liang, Hang Su, Rongdi Yin, Lin Gui, Min Yang, Qin Zhao, Xiaoqi Yu and Ruifeng Xu ..	208
<i>DILBERT: Customized Pre-Training for Domain Adaptation with Category Shift, with an Application to Aspect Extraction</i>	
Entony Lekhtman, Yftah Ziser and Roi Reichart	219
<i>Improving Multimodal fusion via Mutual Dependency Maximisation</i>	
Pierre Colombo, Emile Chapuis, Matthieu Labeau and Chloé Clavel	231
<i>Learning Implicit Sentiment in Aspect-based Sentiment Analysis with Supervised Contrastive Pre-Training</i>	
Zhengyan Li, Yicheng Zou, Chong Zhang, Qi Zhang and Zhongyu Wei	246
<i>Progressive Self-Training with Discriminator for Aspect Term Extraction</i>	
Qianlong Wang, zhiyuan wen, Qin Zhao, Min Yang and Ruifeng Xu	257
<i>Reinforced Counterfactual Data Augmentation for Dual Sentiment Classification</i>	
Hao Chen, Rui Xia and Jianfei Yu	269
<i>Idiosyncratic but not Arbitrary: Learning Idiolects in Online Registers Reveals Distinctive yet Consistent Individual Styles</i>	
Jian Zhu and David Jurgens	279
<i>Narrative Theory for Computational Narrative Understanding</i>	
Andrew Piper, Richard Jean So and David Bamman	298
<i>(Mis)alignment Between Stance Expressed in Social Media Data and Public Opinion Surveys</i>	
Kenneth Joseph, Sarah Shugars, Ryan Gallagher, Jon Green, Alexi Quintana Mathé, Zijian An and David Lazer	312
<i>How Does Counterfactually Augmented Data Impact Models for Social Computing Constructs?</i>	
Indira Sen, Mattia Samory, Fabian Flöck, Claudia Wagner and Isabelle Augenstein	325
<i>Latent Hatred: A Benchmark for Understanding Implicit Hate Speech</i>	
Mai ElSherief, Caleb Ziemis, David Muchlinski, Vaishnavi Anupindi, Jordyn Seybolt, Munmun De Choudhury and Diyi Yang	345
<i>Distilling Linguistic Context for Language Model Compression</i>	
Geondo Park, Gyeongman Kim and Eunho Yang	364
<i>Dynamic Knowledge Distillation for Pre-trained Language Models</i>	
Lei Li, Yankai Lin, Shuhuai Ren, Peng Li, Jie Zhou and Xu Sun	379
<i>Few-Shot Text Generation with Natural Language Instructions</i>	
Timo Schick and Hinrich Schütze	390
<i>SOM-NCSCM: An Efficient Neural Chinese Sentence Compression Model Enhanced with Self-Organizing Map</i>	
Kangli Zi, Shi Wang, Yu Liu, Jicun Li, Yanan Cao and Cungen Cao	403

<i>Efficient Multi-Task Auxiliary Learning: Selecting Auxiliary Data by Feature Similarity</i> Po-Nien Kung, Sheng-Siang Yin, Yi-Cheng Chen, Tse-Hsuan Yang and Yun-Nung Chen	416
<i>GOLD: Improving Out-of-Scope Detection in Dialogues using Data Augmentation</i> Derek Chen and Zhou Yu	429
<i>Graph Based Network with Contextualized Representations of Turns in Dialogue</i> Bongseok Lee and Yong Suk Choi	443
<i>Automatically Exposing Problems with Neural Dialog Models</i> Dian Yu and Kenji Sagae	456
<i>Event Coreference Data (Almost) for Free: Mining Hyperlinks from Online News</i> Michael Bugert and Iryna Gurevych	471
<i>Inducing Stereotypical Character Roles from Plot Structure</i> Labiba Jahan, Rahul Mittal and Mark Finlayson	492
<i>Multitask Semi-Supervised Learning for Class-Imbalanced Discourse Classification</i> Alexander Spangher, Jonathan May, Sz-Rung Shiang and Lingjia Deng	498
<i>Low Frequency Names Exhibit Bias and Overfitting in Contextualizing Language Models</i> Robert Wolfe and Aylin Caliskan	518
<i>Mitigating Language-Dependent Ethnic Bias in BERT</i> Jaimeen Ahn and Alice Oh	533
<i>Adversarial Scrubbing of Demographic Information for Text Classification</i> Somnath Basu Roy Chowdhury, Sayan Ghosh, Yiyuan Li, Junier Oliva, Shashank Srivastava and Snigdha Chaturvedi	550
<i>Open-domain clarification question generation without question examples</i> Julia White, Gabriel Poesia, Robert Hawkins, Dorsa Sadigh and Noah Goodman	563
<i>Improving Sequence-to-Sequence Pre-training via Sequence Span Rewriting</i> Wangchunshu Zhou, Tao Ge, Canwen Xu, Ke Xu and Furu Wei	571
<i>Coarse2Fine: Fine-grained Text Classification on Coarsely-grained Annotated Data</i> Dheeraj Mekala, Varun Gangal and Jingbo Shang	583
<i>Text2Mol: Cross-Modal Molecule Retrieval with Natural Language Queries</i> Carl Edwards, ChengXiang Zhai and Heng Ji	595
<i>Classification of hierarchical text using geometric deep learning: the case of clinical trials corpus</i> Sohrab Ferdowsi, Nikolay Borissov, Julien Knafou, Poorya Amini and Douglas Teodoro	608
<i>The Devil is in the Detail: Simple Tricks Improve Systematic Generalization of Transformers</i> Róbert Csordás, Kazuki Irie and Juergen Schmidhuber	619
<i>Artificial Text Detection via Examining the Topology of Attention Maps</i> Laida Kushnareva, Daniil Cherniavskii, Vladislav Mikhailov, Ekaterina Artemova, Serguei Baranikov, Alexander Bernstein, Irina Piontkovskaya, Dmitri Piontkovski and Evgeny Burnaev	635
<i>Active Learning by Acquiring Contrastive Examples</i> Katerina Margatina, Giorgos Vernikos, Loïc Barrault and Nikolaos Aletras	650

<i>Conditional Poisson Stochastic Beams</i>	
Clara Meister, Afra Amini, Tim Vieira and Ryan Cotterell	664
<i>Building Adaptive Acceptability Classifiers for Neural NLG</i>	
Soumya Batra, Shashank Jain, Peyman Heidari, Ankit Arun, Catharine Youngs, Xintong Li, Pinar Donmez, Shawn Mei, Shiunzu Kuo, Vikas Bhardwaj, Anuj Kumar and Michael White	682
<i>Moral Stories: Situated Reasoning about Norms, Intents, Actions, and their Consequences</i>	
Denis Emelin, Ronan Le Bras, Jena D. Hwang, Maxwell Forbes and Yejin Choi	698
<i>Truth-Conditional Captions for Time Series Data</i>	
Harsh Jhamtani and Taylor Berg-Kirkpatrick	719
<i>Injecting Entity Types into Entity-Guided Text Generation</i>	
Xiangyu Dong, Wenhao Yu, Chenguang Zhu and Meng Jiang	734
<i>Smelting Gold and Silver for Improved Multilingual AMR-to-Text Generation</i>	
Leonardo F. R. Ribeiro, Jonas Pfeiffer, Yue Zhang and Iryna Gurevych	742
<i>Learning Compact Metrics for MT</i>	
Amy Pu, Hyung Won Chung, Ankur Parikh, Sebastian Gehrmann and Thibault Sellam	751
<i>The Impact of Positional Encodings on Multilingual Compression</i>	
Vinit Ravishankar and Anders Søgaard	763
<i>Disentangling Representations of Text by Masking Transformers</i>	
Xiongyi Zhang, Jan-Willem van de Meent and Byron Wallace	778
<i>Exploring the Role of BERT Token Representations to Explain Sentence Probing Results</i>	
Hosein Mohebbi, Ali Modarressi and Mohammad Taher Pilehvar	792
<i>Do Long-Range Language Models Actually Use Long-Range Context?</i>	
Simeng Sun, Kalpesh Krishna, Andrew Mattarella-Micke and Mohit Iyyer	807
<i>The World of an Octopus: How Reporting Bias Influences a Language Model’s Perception of Color</i>	
Cory Paik, Stéphane Aroca-Ouellette, Alessandro Roncone and Katharina Kann	823
<i>SELFEXPLAIN: A Self-Explaining Architecture for Neural Text Classifiers</i>	
Dheeraj Rajagopal, Vidhisha Balachandran, Eduard H Hovy and Yulia Tsvetkov	836
<i>Memory and Knowledge Augmented Language Models for Inferring Salience in Long-Form Stories</i>	
David Wilmot and Frank Keller	851
<i>Semantic Novelty Detection in Natural Language Descriptions</i>	
Nianzu Ma, Alexander Politowicz, Sahisnu Mazumder, Jiahua Chen, Bing Liu, Eric Robertson and Scott Grigsby	866
<i>Jump-Starting Item Parameters for Adaptive Language Tests</i>	
Arya D. McCarthy, Kevin P. Yancey, Geoff T. LaFlair, Jesse Egbert, Manqian Liao and Burr Settles	883
<i>Voice Query Auto Completion</i>	
Raphael Tang, Karun Kumar, Kendra Chalkley, Ji Xin, Liming Zhang, Wenyan Li, Gefei Yang, Yajie Mao, Junho Shin, Geoffrey Craig Murray and Jimmy Lin	900

<i>CoPHE: A Count-Preserving Hierarchical Evaluation Metric in Large-Scale Multi-Label Text Classification</i>	
Matúš Falis, Hang Dong, Alexandra Birch and Beatrice Alex	907
<i>Learning Universal Authorship Representations</i>	
Rafael A. Rivera-Soto, Olivia Elizabeth Miano, Juanita Ordonez, Barry Y. Chen, Aleem Khan, Marcus Bishop and Nicholas Andrews.....	913
<i>Predicting emergent linguistic compositions through time: Syntactic frame extension via multimodal chaining</i>	
Lei Yu and Yang Xu	920
<i>Frequency Effects on Syntactic Rule Learning in Transformers</i>	
Jason Wei, Dan Garrette, Tal Linzen and Ellie Pavlick.....	932
<i>A surprisal–duration trade-off across and within the world’s languages</i>	
Tiago Pimentel, Clara Meister, Elizabeth Salesky, Simone Teufel, Damián Blasi and Ryan Cotterell	949
<i>Revisiting the Uniform Information Density Hypothesis</i>	
Clara Meister, Tiago Pimentel, Patrick Haller, Lena Jäger, Ryan Cotterell and Roger Levy	963
<i>Condenser: a Pre-training Architecture for Dense Retrieval</i>	
Luyu Gao and Jamie Callan.....	981
<i>Monitoring geometrical properties of word embeddings for detecting the emergence of new topics.</i>	
Clément Christophe, Julien Velcin, Jairo Cugliari, Manel Boumghar and Philippe Suignard ...	994
<i>Contextualized Query Embeddings for Conversational Search</i>	
Sheng-Chieh Lin, Jheng-Hong Yang and Jimmy Lin	1004
<i>Ultra-High Dimensional Sparse Representations with Binarization for Efficient Text Retrieval</i>	
Kyoung-Rok Jang, Junmo Kang, Giwon Hong, Sung-Hyon Myaeng, Joohee Park, Taewon Yoon and Heecheol Seo	1016
<i>IR like a SIR: Sense-enhanced Information Retrieval for Multiple Languages</i>	
Rexhina Blloshmi, Tommaso Pasini, Niccolò Campolungo, Somnath Banerjee, Roberto Navigli and Gabriella Pasi.....	1030
<i>Neural Attention-Aware Hierarchical Topic Model</i>	
YUAN JIN, He Zhao, Ming Liu, Lan Du and Wray Buntine	1042
<i>Relational World Knowledge Representation in Contextual Language Models: A Review</i>	
Tara Safavi and Danai Koutra	1053
<i>Certified Robustness to Programmable Transformations in LSTMs</i>	
Yuhao Zhang, Aws Albarghouthi and Loris D’Antoni	1068
<i>ReGen: Reinforcement Learning for Text and Knowledge Base Generation using Pretrained Language Models</i>	
Pierre Dognin, Inkit Padhi, Igor Melnyk and Payel Das.....	1084
<i>Contrastive Out-of-Distribution Detection for Pretrained Transformers</i>	
Wenxuan Zhou, Fangyu Liu and Muhao Chen	1100

<i>MindCraft: Theory of Mind Modeling for Situated Dialogue in Collaborative Tasks</i> Cristian-Paul Bara, Sky CH-Wang and Joyce Chai	1112
<i>Detecting Speaker Personas from Conversational Texts</i> Jia-Chen Gu, Zhenhua Ling, Yu Wu, Quan Liu, Zhigang Chen and Xiaodan Zhu	1126
<i>Cross-lingual Intermediate Fine-tuning improves Dialogue State Tracking</i> Nikita Moghe, Mark Steedman and Alexandra Birch	1137
<i>ConvFiT: Conversational Fine-Tuning of Pretrained Language Models</i> Ivan Vulić, Pei-Hao Su, Samuel Coope, Daniela Gerz, Paweł Budzianowski, Iñigo Casanueva, Nikola Mrkšić and Tsung-Hsien Wen	1151
<i>We've had this conversation before: A Novel Approach to Measuring Dialog Similarity</i> Ofar Lavi, Ella Rabinovich, Segev Shlomov, David Boaz, Inbal Ronen and Ateret Anaby Tavor	1169
<i>Towards Incremental Transformers: An Empirical Analysis of Transformer Models for Incremental NLU</i> Patrick Kahardipraja, Brielen Madureira and David Schlangen	1178
<i>Feedback Attribution for Counterfactual Bandit Learning in Multi-Domain Spoken Language Understanding</i> Tobias Falke and Patrick Lehnen	1190
<i>Label Verbalization and Entailment for Effective Zero and Few-Shot Relation Extraction</i> Oscar Sainz, Oier Lopez de Lacalle, Gorka Labaka, Ander Barrena and Eneko Agirre	1199
<i>Extend, don't rebuild: Phrasing conditional graph modification as autoregressive sequence labelling</i> Leon Weber, Jannes Münchmeyer, Samuele Garda and Ulf Leser	1213
<i>Zero-Shot Information Extraction as a Unified Text-to-Triple Translation</i> Chenguang Wang, Xiao Liu, Zui Chen, Haoyun Hong, Jie Tang and Dawn Song	1225
<i>Learning Logic Rules for Document-Level Relation Extraction</i> Dongyu Ru, Changzhi Sun, Jiangtao Feng, Lin Qiu, Hao Zhou, Weinan Zhang, Yong Yu and Lei Li	1239
<i>A Large-Scale Dataset for Empathetic Response Generation</i> Anuradha Welivita, Yubo Xie and Pearl Pu	1251
<i>The Perils of Using Mechanical Turk to Evaluate Open-Ended Text Generation</i> Marzena Karpinska, Nader Akoury and Mohit Iyyer	1265
<i>Documenting Large Webtext Corpora: A Case Study on the Colossal Clean Crawled Corpus</i> Jesse Dodge, Maarten Sap, Ana Marasović, William Agnew, Gabriel Ilharco, Dirk Groeneveld, Margaret Mitchell and Matt Gardner	1286
<i>AfroMT: Pretraining Strategies and Reproducible Benchmarks for Translation of 8 African Languages</i> Machel Reid, Junjie Hu, Graham Neubig and Yutaka Matsuo	1306
<i>Evaluating the Evaluation Metrics for Style Transfer: A Case Study in Multilingual Formality Transfer</i> Eleftheria Briakou, Sweta Agrawal, Joel Tetreault and Marine Carpuat	1321
<i>MS-Mentions: Consistently Annotating Entity Mentions in Materials Science Procedural Text</i> Tim O'Gorman, Zach Jensen, Sheshera Mysore, Kevin Huang, Rubayyat Mahbub, Elsa Olivetti and Andrew McCallum	1337

<i>Understanding Politics via Contextualized Discourse Processing</i>	
Rajkumar Pujari and Dan Goldwasser	1353
<i>Conundrums in Event Coreference Resolution: Making Sense of the State of the Art</i>	
Jing Lu and Vincent Ng	1368
<i>Weakly supervised discourse segmentation for multiparty oral conversations</i>	
Lila Gravelier, Julie Hunter, Philippe Muller, Thomas Pellegrini and Isabelle Ferrané	1381
<i>Narrative Embedding: Re-Contextualization Through Attention</i>	
Sean Wilner, Daniel Woolridge and Madeleine Glick	1393
<i>Focus on what matters: Applying Discourse Coherence Theory to Cross Document Coreference</i>	
William Held, Dan Iter and Dan Jurafsky	1406
<i>Salience-Aware Event Chain Modeling for Narrative Understanding</i>	
Xiyang Zhang, Muhao Chen and Jonathan May	1418
<i>Asking It All: Generating Contextualized Questions for any Semantic Role</i>	
Valentina Pyatkin, Paul Roit, Julian Michael, Yoav Goldberg, Reut Tsarfaty and Ido Dagan ..	1429
<i>Fast, Effective, and Self-Supervised: Transforming Masked Language Models into Universal Lexical and Sentence Encoders</i>	
Fangyu Liu, Ivan Vulić, Anna Korhonen and Nigel Collier	1442
<i>RuleBERT: Teaching Soft Rules to Pre-Trained Language Models</i>	
Mohammed Saeed, Naser Ahmadi, Preslav Nakov and Paolo Papotti	1460
<i>Stepmothers are mean and academics are pretentious: What do pretrained language models learn about you?</i>	
Rochelle Choenni, Ekaterina Shutova and Robert van Rooij	1477
<i>ConSeC: Word Sense Disambiguation as Continuous Sense Comprehension</i>	
Edoardo Barba, Luigi Procopio and Roberto Navigli	1492
<i>Shortcutted Commonsense: Data Spuriousness in Deep Learning of Commonsense Reasoning</i>	
Ruben Branco, António Branco, João António Rodrigues and João Ricardo Silva	1504
<i>When differential privacy meets NLP: The devil is in the detail</i>	
Ivan Habernal	1522
<i>Achieving Model Robustness through Discrete Adversarial Training</i>	
Maor Ivgi and Jonathan Berant	1529
<i>Debiasing Methods in Natural Language Understanding Make Bias More Accessible</i>	
Michael Mendelson and Yonatan Belinkov	1545
<i>Evaluating the Robustness of Neural Language Models to Input Perturbations</i>	
Milad Moradi and Matthias Samwald	1558
<i>How much pretraining data do language models need to learn syntax?</i>	
Laura Pérez-Mayos, Miguel Ballesteros and Leo Wanner	1571
<i>Sorting through the noise: Testing robustness of information processing in pre-trained language models</i>	
Lalchand Pandia and Allyson Ettinger	1583

<i>Contrastive Explanations for Model Interpretability</i>	
Alon Jacovi, Swabha Swayamdipta, Shauli Ravfogel, Yanai Elazar, Yejin Choi and Yoav Goldberg	
1597	
<i>On the Transferability of Adversarial Attacks against Neural Text Classifier</i>	
Liping Yuan, Xiaoqing Zheng, Yi Zhou, Cho-Jui Hsieh and Kai-Wei Chang	1612
<i>Conditional probing: measuring usable information beyond a baseline</i>	
John Hewitt, Kawin Ethayarajh, Percy Liang and Christopher Manning	1626
<i>GFST: Gender-Filtered Self-Training for More Accurate Gender in Translation</i>	
Prafulla Kumar Choubey, Anna Currey, Prashant Mathur and Georgiana Dinu	1640
<i>"Wikily" Supervised Neural Translation Tailored to Cross-Lingual Tasks</i>	
Mohammad Sadegh Rasooli, Chris Callison-Burch and Derry Tanti Wijaya	1655
<i>mT6: Multilingual Pretrained Text-to-Text Transformer with Translation Pairs</i>	
Zewen Chi, Li Dong, Shuming Ma, Shaohan Huang, Saksham Singhal, Xian-Ling Mao, Heyan Huang, Xia Song and Furu Wei	1671
<i>Improving Zero-Shot Cross-Lingual Transfer Learning via Robust Training</i>	
Kuan-Hao Huang, Wasi Ahmad, Nanyun Peng and Kai-Wei Chang	1684
<i>Speechformer: Reducing Information Loss in Direct Speech Translation</i>	
Sara Papi, Marco Gaido, Matteo Negri and Marco Turchi	1698
<i>Is "moby dick" a Whale or a Bird? Named Entities and Terminology in Speech Translation</i>	
Marco Gaido, Susana Rodríguez, Matteo Negri, Luisa Bentivogli and Marco Turchi	1707
<i>HintedBT: Augmenting Back-Translation with Quality and Transliteration Hints</i>	
Sahana Ramnath, Melvin Johnson, Abhirut Gupta and Aravindan Raghuv eer	1717
<i>Translation-based Supervision for Policy Generation in Simultaneous Neural Machine Translation</i>	
Ashkan Alinejad, Hassan S. Shavarani and Anoop Sarkar	1734
<i>Nearest Neighbour Few-Shot Learning for Cross-lingual Classification</i>	
M Saiful Bari, Batool Haider and Saab Mansour	1745
<i>Cross-Attention is All You Need: Adapting Pretrained Transformers for Machine Translation</i>	
Mozhdeh Gheini, Xiang Ren and Jonathan May	1754
<i>Effects of Parameter Norm Growth During Transformer Training: Inductive Bias from Gradient Descent</i>	
William Merrill, Vivek Ramanujan, Yoav Goldberg, Roy Schwartz and Noah A. Smith	1766
<i>Foreseeing the Benefits of Incidental Supervision</i>	
Hangfeng He, Mingyuan Zhang, Qiang Ning and Dan Roth	1782
<i>Competency Problems: On Finding and Removing Artifacts in Language Data</i>	
Matt Gardner, William Merrill, Jesse Dodge, Matthew Peters, Alexis Ross, Sameer Singh and Noah A. Smith	1801
<i>Knowledge-Aware Meta-learning for Low-Resource Text Classification</i>	
Huaxiu Yao, Ying-xin Wu, Maruan Al-Shedivat and Eric Xing	1814
<i>Sentence Bottleneck Autoencoders from Transformer Language Models</i>	
Ivan Montero, Nikolaos Pappas and Noah A. Smith	1822

<i>Efficient Contrastive Learning via Novel Data Augmentation and Curriculum Learning</i> Seonghyeon Ye, Jiseon Kim and Alice Oh	1832
<i>CR-Walker: Tree-Structured Graph Reasoning and Dialog Acts for Conversational Recommendation</i> Wenchang Ma, Ryuichi Takanobu and Minlie Huang	1839
<i>DIALKI: Knowledge Identification in Conversational Systems through Dialogue-Document Contextualization</i> Zequ Wu, Bo-Ru Lu, Hannaneh Hajishirzi and Mari Ostendorf	1852
<i>Iconary: A Pictionary-Based Game for Testing Multimodal Communication with Drawings and Text</i> Christopher Clark, Jordi Salvador, Dustin Schwenk, Derrick Bonafilia, Mark Yatskar, Eric Kolve, Alvaro Herrasti, Jonghyun Choi, Sachin Mehta, Sam Skjonsberg, Carissa Schoenick, Aaron Sarnat, Hannaneh Hajishirzi, Aniruddha Kembhavi, Oren Etzioni and Ali Farhadi	1864
<i>Self-training Improves Pre-training for Few-shot Learning in Task-oriented Dialog Systems</i> Fei Mi, Wanhao Zhou, Lingjing Kong, Fengyu Cai, Minlie Huang and Boi Faltings	1887
<i>Contextual Rephrase Detection for Reducing Friction in Dialogue Systems</i> Zhuoyi Wang, Saurabh Gupta, Jie Hao, Xing Fan, Dingcheng Li, Alexander Hanbo Li and Chenlei Guo	1899
<i>Few-Shot Intent Detection via Contrastive Pre-Training and Fine-Tuning</i> Jianguo Zhang, Trung Bui, Seunghyun Yoon, Xiang Chen, Zhiwei Liu, Congying Xia, Quan Hung Tran, Walter Chang and Philip Yu	1906
<i>"It doesn't look good for a date": Transforming Critiques into Preferences for Conversational Recommendation Systems</i> Victor Bursztyrn, Jennifer Healey, Nedim Lipka, Eunye Koh, Doug Downey and Larry Birnbaum	1913
<i>AttentionRank: Unsupervised Keyphrase Extraction using Self and Cross Attentions</i> Haoran Ding and Xiao Luo	1919
<i>Unsupervised Relation Extraction: A Variational Autoencoder Approach</i> Chenhan Yuan and Hoda Eldardiry	1929
<i>Robust Retrieval Augmented Generation for Zero-shot Slot Filling</i> Michael Glass, Gaetano Rossiello, Md Faisal Mahbub Chowdhury and Alfio Gliozzo	1939
<i>Everything Is All It Takes: A Multipronged Strategy for Zero-Shot Cross-Lingual Information Extraction</i> Mahsa Yarmohammadi, Shijie Wu, Marc Marone, Haoran Xu, Seth Ebner, Guanghui Qin, Yunmo Chen, Jialiang Guo, Craig Harman, Kenton Murray, Aaron Steven White, Mark Dredze and Benjamin Van Durme	1950
<i>Harms of Gender Exclusivity and Challenges in Non-Binary Representation in Language Technologies</i> Sunipa Dev, Masoud Monajatipoor, Anaelia Ovalle, Arjun Subramonian, Jeff Phillips and Kai-Wei Chang	1968
<i>Are Gender-Neutral Queries Really Gender-Neutral? Mitigating Gender Bias in Image Search</i> Jialu Wang, Yang Liu and Xin Wang	1995
<i>Style Pooling: Automatic Text Style Obfuscation for Improved Classification Fairness</i> Fatemehsadat Miresghallah and Taylor Berg-Kirkpatrick	2009

<i>Modeling Disclosive Transparency in NLP Application Descriptions</i>	
Michael Saxon, Sharon Levy, Xinyi Wang, Alon Albalak and William Yang Wang	2023
<i>Reconstruction Attack on Instance Encoding for Language Understanding</i>	
Shangyu Xie and Yuan Hong	2038
<i>Fairness-aware Class Imbalanced Learning</i>	
Shivashankar Subramanian, Afshin Rahimi, Timothy Baldwin, Trevor Cohn and Lea Frermann	2045
<i>CRYPTOGRU: Low Latency Privacy-Preserving Text Analysis With GRU</i>	
Bo Feng, Qian Lou, Lei Jiang and Geoffrey Fox	2052
<i>Local Word Discovery for Interactive Transcription</i>	
William Lane and Steven Bird	2058
<i>Segment, Mask, and Predict: Augmenting Chinese Word Segmentation with Self-Supervision</i>	
Mieradilijiang Maimaiti, Yang Liu, Yuanhang Zheng, Gang Chen, Kaiyu Huang, Ji Zhang, Huanbo Luan and Maosong Sun	2068
<i>Minimal Supervision for Morphological Inflection</i>	
Omer Goldman and Reut Tsarfaty	2078
<i>Fast WordPiece Tokenization</i>	
Xinying Song, Alex Salcianu, Yang Song, Dave Dopson and Denny Zhou	2089
<i>You should evaluate your language model on marginal likelihood over tokenisations</i>	
Kris Cao and Laura Rimell	2104
<i>Broaden the Vision: Geo-Diverse Visual Commonsense Reasoning</i>	
Da Yin, Liunian Harold Li, Ziniu Hu, Nanyun Peng and Kai-Wei Chang	2115
<i>Reference-Centric Models for Grounded Collaborative Dialogue</i>	
Daniel Fried, Justin Chiu and Dan Klein	2130
<i>CrossVQA: Scalably Generating Benchmarks for Systematically Testing VQA Generalization</i>	
Arjun Akula, Soravit Changpinyo, Boqing Gong, Piyush Sharma, Song-Chun Zhu and Radu Soricut	2148
<i>Visual Goal-Step Inference using wikiHow</i>	
Yue Yang, Artemis Panagopoulou, Qing Lyu, Li Zhang, Mark Yatskar and Chris Callison-Burch	2167
<i>Systematic Generalization on gSCAN: What is Nearly Solved and What is Next?</i>	
Linlu Qiu, Hexiang Hu, Bowen Zhang, Peter Shaw and Fei Sha	2180
<i>Effect of Visual Extensions on Natural Language Understanding in Vision-and-Language Models</i>	
Taichi Iki and Akiko Aizawa	2189
<i>Neural Path Hunter: Reducing Hallucination in Dialogue Systems via Path Grounding</i>	
Nouha Dziri, Andrea Madotto, Osmar Zaiane and Avishek Joey Bose	2197
<i>Thinking Clearly, Talking Fast: Concept-Guided Non-Autoregressive Generation for Open-Domain Dialogue Systems</i>	
Yicheng Zou, Zhihua Liu, Xingwu Hu and Qi Zhang	2215

<i>Perspective-taking and Pragmatics for Generating Empathetic Responses Focused on Emotion Causes</i> Hyunwoo Kim, Byeongchang Kim and Gunhee Kim	2227
<i>Generation and Extraction Combined Dialogue State Tracking with Hierarchical Ontology Integration</i> Xinmeng Li, Qian Li, Wansen Wu and Quanjun Yin	2241
<i>CoLV: A Collaborative Latent Variable Model for Knowledge-Grounded Dialogue Generation</i> Haolan Zhan, Lei Shen, Hongshen Chen and Hainan Zhang	2250
<i>A Three-Stage Learning Framework for Low-Resource Knowledge-Grounded Dialogue Generation</i> Shilei Liu, Xiaofeng Zhao, Bochao Li, Feiliang Ren, Longhui Zhang and Shujuan Yin	2262
<i>Intention Reasoning Network for Multi-Domain End-to-end Task-Oriented Dialogue</i> Zhiyuan Ma, Jianjun Li, Zezheng Zhang, Guohui Li and Yongjing Cheng	2273
<i>More is Better: Enhancing Open-Domain Dialogue Generation via Multi-Source Heterogeneous Knowledge</i> Sixing Wu, Ying Li, Minghui Wang, Dawei Zhang, Yang Zhou and Zhonghai Wu	2286
<i>Domain-Lifelong Learning for Dialogue State Tracking via Knowledge Preservation Networks</i> Qingbin Liu, Pengfei Cao, Cao Liu, Jiansong Chen, Xunliang Cai, Fan Yang, Shizhu He, Kang Liu and Jun Zhao	2301
<i>CSAGN: Conversational Structure Aware Graph Network for Conversational Semantic Role Labeling</i> Han Wu, Kun Xu and Linqi Song	2312
<i>Different Strokes for Different Folks: Investigating Appropriate Further Pre-training Approaches for Diverse Dialogue Tasks</i> Yao Qiu, Jinchao Zhang and Jie Zhou	2318
<i>Knowledge Enhanced Fine-Tuning for Better Handling Unseen Entities in Dialogue Generation</i> Leyang Cui, Yu Wu, Shujie Liu and Yue Zhang	2328
<i>An Evaluation Dataset and Strategy for Building Robust Multi-turn Response Selection Model</i> Kijong Han, Seojin Lee and Dong-hun Lee	2338
<i>Unsupervised Conversation Disentanglement through Co-Training</i> Hui Liu, Zhan Shi and Xiaodan Zhu	2345
<i>Don't be Contradicted with Anything! CI-ToD: Towards Benchmarking Consistency for Task-oriented Dialogue System</i> Libo Qin, Tianbao Xie, Shijue Huang, Qiguang Chen, Xiao Xu and Wanxiang Che	2357
<i>Transferable Persona-Grounded Dialogues via Grounded Minimal Edits</i> Chen Henry Wu, Yinhe Zheng, Xiaoxi Mao and Minlie Huang	2368
<i>EARL: Informative Knowledge-Grounded Conversation Generation with Entity-Agnostic Representation Learning</i> Hao Zhou, Minlie Huang, Yong Liu, Wei Chen and Xiaoyan Zhu	2383
<i>DialogueCSE: Dialogue-based Contrastive Learning of Sentence Embeddings</i> Che Liu, Rui Wang, Jinghua Liu, Jian Sun, Fei Huang and Luo Si	2396
<i>Improving Graph-based Sentence Ordering with Iteratively Predicted Pairwise Orderings</i> Shaopeng Lai, Ante Wang, Fandong Meng, Jie Zhou, Yubin Ge, Jiali Zeng, Junfeng Yao, Degen Huang and Jinsong Su	2407

<i>Not Just Classification: Recognizing Implicit Discourse Relation on Joint Modeling of Classification and Generation</i>	
Feng Jiang, Yaxin Fan, Xiaomin Chu, Peifeng Li and Qiaoming Zhu	2418
<i>A Language Model-based Generative Classifier for Sentence-level Discourse Parsing</i>	
Ying Zhang, Hidetaka Kamigaito and Manabu Okumura	2432
<i>Multimodal Phased Transformer for Sentiment Analysis</i>	
Junyan Cheng, Iordanis Fostiropoulos, Barry Boehm and Mohammad Soleymani	2447
<i>Hierarchical Multi-label Text Classification with Horizontal and Vertical Category Correlations</i>	
Linli Xu, Sijie Teng, Ruoyu Zhao, Junliang Guo, Chi Xiao, Deqiang Jiang and Bo Ren	2459
<i>RankNAS: Efficient Neural Architecture Search by Pairwise Ranking</i>	
Chi Hu, Chenglong Wang, Xiangnan Ma, Xia Meng, Yinqiao Li, Tong Xiao, Jingbo Zhu and changliang li	2469
<i>FLiText: A Faster and Lighter Semi-Supervised Text Classification with Convolution Networks</i>	
Chen Liu, zhang mengchao, Fu Zhibing, Panpan Hou and Yu Li	2481
<i>Evaluating Debiasing Techniques for Intersectional Biases</i>	
Shivashankar Subramanian, Xudong Han, Timothy Baldwin, Trevor Cohn and Lea Frermann	2492
<i>Definition Modelling for Appropriate Specificity</i>	
Han Huang, Tomoyuki Kajiwara and Yuki Arase	2499
<i>Transductive Learning for Unsupervised Text Style Transfer</i>	
Fei Xiao, Liang Pang, Yanyan Lan, Yan Wang, Huawei Shen and Xueqi Cheng	2510
<i>Integrating Semantic Scenario and Word Relations for Abstractive Sentence Summarization</i>	
Yong Guan, Shaoru Guo, Ru Li, Xiaoli Li and Hu Zhang	2522
<i>Coupling Context Modeling with Zero Pronoun Recovering for Document-Level Natural Language Generation</i>	
Xin Tan, Longyin Zhang and Guodong Zhou	2530
<i>Adaptive Bridge between Training and Inference for Dialogue Generation</i>	
Haoran Xu, Hainan Zhang, Yanyan Zou, Hongshen Chen, Zhuoye Ding and Yanyan Lan	2541
<i>ConRPG: Paraphrase Generation using Contexts as Regularizer</i>	
Yuxian Meng, Xiang Ao, Qing He, Xiaofei Sun, Qinghong Han, Fei Wu, Chun Fan and Jiwei Li	2551
<i>Building the Directed Semantic Graph for Coherent Long Text Generation</i>	
Ziao Wang, Xiaofeng Zhang and Hongwei Du	2563
<i>Iterative GNN-based Decoder for Question Generation</i>	
Zichu Fei, Qi Zhang and Yaqian Zhou	2573
<i>Asking Questions Like Educational Experts: Automatically Generating Question-Answer Pairs on Real-World Examination Data</i>	
Fanyi Qu, Xin Jia and Yunfang Wu	2583
<i>Syntactically-Informed Unsupervised Paraphrasing with Non-Parallel Data</i>	
Erguang Yang, Mingtong Liu, Deyi Xiong, Yujie Zhang, Yao Meng, Changjian Hu, Jinan Xu and Yufeng Chen	2594

<i>Exploring Task Difficulty for Few-Shot Relation Extraction</i>	
Jiale Han, Bo Cheng and Wei Lu	2605
<i>MuVER: Improving First-Stage Entity Retrieval with Multi-View Entity Representations</i>	
Xinyin Ma, Yong Jiang, NGUYEN BACH, Tao Wang, Zhongqiang Huang, Fei Huang and Weiming Lu	2617
<i>Treasures Outside Contexts: Improving Event Detection via Global Statistics</i>	
Rui Li, Wenlin Zhao, Cheng Yang and Sen Su	2625
<i>Uncertain Local-to-Global Networks for Document-Level Event Factuality Identification</i>	
Pengfei Cao, Yubo Chen, Yuqing Yang, Kang Liu and Jun Zhao	2636
<i>A Novel Global Feature-Oriented Relational Triple Extraction Model based on Table Filling</i>	
Feiliang Ren, Longhui Zhang, Shujuan Yin, Xiaofeng Zhao, Shilei Liu, Bochao Li and Yaduo Liu	2646
<i>Structure-Augmented Keyphrase Generation</i>	
Jihyuk Kim, Myeongho Jeong, Seungtaek Choi and Seung-won Hwang	2657
<i>An Empirical Study on Multiple Information Sources for Zero-Shot Fine-Grained Entity Typing</i>	
Yi Chen, Haiyun Jiang, Lema Liu, Shuming Shi, Chuang Fan, Min Yang and Ruifeng Xu ..	2668
<i>DyLex: Incorporating Dynamic Lexicons into BERT for Sequence Labeling</i>	
Baojun Wang, Zhao Zhang, Kun Xu, Guang-Yuan Hao, Yuyang Zhang, Lifeng Shang, Linlin Li, Xiao Chen, Xin Jiang and Qun Liu	2679
<i>MapRE: An Effective Semantic Mapping Approach for Low-resource Relation Extraction</i>	
Manqing Dong, Chunguang Pan and Zhipeng Luo	2694
<i>Heterogeneous Graph Neural Networks for Keyphrase Generation</i>	
Jiacheng Ye, Ruijian Cai, Tao Gui and Qi Zhang	2705
<i>Machine Reading Comprehension as Data Augmentation: A Case Study on Implicit Event Argument Extraction</i>	
Jian Liu, Yufeng Chen and Jinan Xu	2716
<i>Importance Estimation from Multiple Perspectives for Keyphrase Extraction</i>	
Mingyang Song, Liping Jing and Lin Xiao	2726
<i>Gradient Imitation Reinforcement Learning for Low Resource Relation Extraction</i>	
Xuming Hu, Chenwei Zhang, Yawen Yang, Xiaohe Li, Li Lin, Lijie Wen and Philip S. Yu ...	2737
<i>Low-resource Taxonomy Enrichment with Pretrained Language Models</i>	
Kunihiro Takeoka, Kosuke Akimoto and Masafumi Oyamada	2747
<i>Entity Relation Extraction as Dependency Parsing in Visually Rich Documents</i>	
Yue Zhang, Zhang Bo, Rui Wang, Junjie Cao, Chen Li and Zuyi Bao	2759
<i>Synchronous Dual Network with Cross-Type Attention for Joint Entity and Relation Extraction</i>	
Hui Wu and xiaodong shi	2769
<i>Less is More: Pretrain a Strong Siamese Encoder for Dense Text Retrieval Using a Weak Decoder</i>	
Shuqi Lu, Di He, Chenyan Xiong, Guolin Ke, Waleed Malik, Zhicheng Dou, Paul Bennett, Tie-Yan Liu and Arnold Overwijk	2780

<i>TransPrompt: Towards an Automatic Transferable Prompting Framework for Few-shot Text Classification</i>	
Chengyu Wang, Jianing Wang, Minghui Qiu, Jun Huang and Ming Gao	2792
<i>Weakly-supervised Text Classification Based on Keyword Graph</i>	
Lu Zhang, Jiandong Ding, Yi Xu, Yingyao Liu and Shuigeng Zhou	2803
<i>Efficient-FedRec: Efficient Federated Learning Framework for Privacy-Preserving News Recommendation</i>	
Jingwei Yi, Fangzhao Wu, Chuhan Wu, Ruixuan Liu, Guangzhong Sun and Xing Xie	2814
<i>RocketQAv2: A Joint Training Method for Dense Passage Retrieval and Passage Re-ranking</i>	
Ruiyang Ren, Yingqi Qu, Jing Liu, Wayne Xin Zhao, QiaoQiao She, Hua Wu, Haifeng Wang and Ji-Rong Wen	2825
<i>Dealing with Typos for BERT-based Passage Retrieval and Ranking</i>	
Shengyao Zhuang and Guido Zuccon	2836
<i>From Alignment to Assignment: Frustratingly Simple Unsupervised Entity Alignment</i>	
Xin Mao, wenting wang, Yuanbin Wu and Man Lan	2843
<i>Simple and Effective Unsupervised Redundancy Elimination to Compress Dense Vectors for Passage Retrieval</i>	
Xueguang Ma, Minghan Li, Kai Sun, Ji Xin and Jimmy Lin	2854
<i>Relation Extraction with Word Graphs from N-grams</i>	
Han Qin, Yuanhe Tian and Yan Song	2860
<i>A Bayesian Framework for Information-Theoretic Probing</i>	
Tiago Pimentel and Ryan Cotterell	2869
<i>Masked Language Modeling and the Distributional Hypothesis: Order Word Matters Pre-training for Little</i>	
Koustuv Sinha, Robin Jia, Dieuwke Hupkes, Joelle Pineau, Adina Williams and Douwe Kiela	2888
<i>What's Hidden in a One-layer Randomly Weighted Transformer?</i>	
Sheng Shen, Zhewei Yao, Douwe Kiela, Kurt Keutzer and Michael Mahoney	2914
<i>Rethinking Denoised Auto-Encoding in Language Pre-Training</i>	
Fuli Luo, Pengcheng Yang, Shicheng Li, Xuancheng Ren, Xu Sun, Songfang Huang and Fei Huang	2922
<i>Lifelong Explainer for Lifelong Learners</i>	
Xuelin Situ, Sameen Maruf, Ingrid Zukerman, Cecile Paris and Gholamreza Haffari	2933
<i>Linguistic Dependencies and Statistical Dependence</i>	
Jacob Louis Hoover, Wenyu Du, Alessandro Sordani and Timothy J. O'Donnell	2941
<i>Modeling Human Sentence Processing with Left-Corner Recurrent Neural Network Grammars</i>	
Ryo Yoshida, Hiroshi Noji and Yohei Oseki	2964
<i>A Simple and Effective Positional Encoding for Transformers</i>	
Pu-Chin Chen, Henry Tsai, Srinadh Bhojanapalli, Hyung Won Chung, Yin-Wen Chang and Chun-Sung Ferng	2974

<i>Explore Better Relative Position Embeddings from Encoding Perspective for Transformer Models</i> Anlin Qu, Jianwei Niu and Shasha Mo	2989
<i>Adversarial Mixing Policy for Relaxing Locally Linear Constraints in Mixup</i> Guang Liu, Yuzhao Mao, Huang Hailong, Gao Weiguo and Li Xuan	2998
<i>Is this the end of the gold standard? A straightforward reference-less grammatical error correction metric</i> Md Asadul Islam and Enrico Magnani	3009
<i>Augmenting BERT-style Models with Predictive Coding to Improve Discourse-level Representations</i> Vladimir Araujo, Andrés Villa, Marcelo Mendoza, Marie-Francine Moens and Alvaro Soto ..	3016
<i>Backdoor Attacks on Pre-trained Models by Layerwise Weight Poisoning</i> Linyang Li, Demin Song, Xiaonan Li, Jiehang Zeng, Ruotian Ma and Xipeng Qiu	3023
<i>GAML-BERT: Improving BERT Early Exiting by Gradient Aligned Mutual Learning</i> Wei Zhu, Xiaoling Wang, Yuan Ni and Guotong Xie	3033
<i>The Power of Scale for Parameter-Efficient Prompt Tuning</i> Brian Lester, Rami Al-Rfou and Noah Constant	3045
<i>Scalable Font Reconstruction with Dual Latent Manifolds</i> Nikita Srivatsan, Si Wu, Jonathan Barron and Taylor Berg-Kirkpatrick	3060
<i>Neuro-Symbolic Approaches for Text-Based Policy Learning</i> Subhajit Chaudhury, Prithviraj Sen, Masaki Ono, Daiki Kimura, Michiaki Tatsubori and Asim Munawar	3073
<i>Layer-wise Model Pruning based on Mutual Information</i> Chun Fan, Jiwei Li, Tianwei Zhang, Xiang Ao, Fei Wu, Yuxian Meng and Xiaofei Sun	3079
<i>Hierarchical Heterogeneous Graph Representation Learning for Short Text Classification</i> Yaqing Wang, Song Wang, Quanming Yao and Dejing Dou	3091
<i>kFolden: k-Fold Ensemble for Out-Of-Distribution Detection</i> Xiaoya Li, Jiwei Li, Xiaofei Sun, Chun Fan, Tianwei Zhang, Fei Wu, Yuxian Meng and Jun Zhang	3102
<i>Frustratingly Simple Pretraining Alternatives to Masked Language Modeling</i> Atsuki Yamaguchi, George Chrysostomou, Katerina Margatina and Nikolaos Aletras	3116
<i>HRKD: Hierarchical Relational Knowledge Distillation for Cross-domain Language Model Compression</i> Chenhe Dong, Yaliang Li, Ying Shen and Minghui Qiu	3126
<i>Searching for an Effective Defender: Benchmarking Defense against Adversarial Word Substitution</i> Zongyi Li, Jianhan Xu, Jiehang Zeng, Linyang Li, Xiaoqing Zheng, Qi Zhang, Kai-Wei Chang and Cho-Jui Hsieh	3137
<i>Re-embedding Difficult Samples via Mutual Information Constrained Semantically Oversampling for Imbalanced Text Classification</i> Jiachen Tian, Shizhan Chen, Xiaowang Zhang, Zhiyong Feng, Deyi Xiong, Shaojuan Wu and Chunliu Dou	3148

<i>Beyond Text: Incorporating Metadata and Label Structure for Multi-Label Document Classification using Heterogeneous Graphs</i>	
Chenchen Ye, Linhai Zhang, Yulan He, Deyu Zhou and Jie Wu	3162
<i>Natural Language Processing Meets Quantum Physics: A Survey and Categorization</i>	
Sixuan Wu, Jian Li, Peng Zhang and Yue Zhang	3172
<i>MetaTS: Meta Teacher-Student Network for Multilingual Sequence Labeling with Minimal Supervision</i>	
Zheng Li, Danqing Zhang, Tianyu Cao, Ying Wei, Yiwei Song and Bing Yin	3183
<i>Neural Machine Translation with Heterogeneous Topic Knowledge Embeddings</i>	
Weixuan Wang, Wei Peng, Meng Zhang and Qun Liu	3197
<i>Allocating Large Vocabulary Capacity for Cross-Lingual Language Model Pre-Training</i>	
Bo Zheng, Li Dong, Shaohan Huang, Saksham Singhal, Wanxiang Che, Ting Liu, Xia Song and Furu Wei	3203
<i>Recurrent Attention for Neural Machine Translation</i>	
Jiali Zeng, Shuangzhi Wu, Yongjing Yin, Yufan Jiang and Mu Li	3216
<i>Learning from Multiple Noisy Augmented Data Sets for Better Cross-Lingual Spoken Language Understanding</i>	
Yingmei Guo, Linjun Shou, Jian Pei, Ming Gong, Mingxing Xu, Zhiyong Wu and Daxin Jiang	3226
<i>Enlivening Redundant Heads in Multi-head Self-attention for Machine Translation</i>	
Tianfu Zhang, Heyan Huang, Chong Feng and Longbing Cao	3238
<i>Unsupervised Neural Machine Translation with Universal Grammar</i>	
Zuchao Li, Masao Utiyama, Eiichiro Sumita and Hai Zhao	3249
<i>Encouraging Lexical Translation Consistency for Document-Level Neural Machine Translation</i>	
Xinglin Lyu, Junhui Li, Zhengxian Gong and Min Zhang	3265
<i>Improving Neural Machine Translation by Bidirectional Training</i>	
Liang Ding, Di Wu and Dacheng Tao	3278
<i>Scheduled Sampling Based on Decoding Steps for Neural Machine Translation</i>	
Yijin Liu, Fandong Meng, Yufeng Chen, Jinan Xu and Jie Zhou	3285
<i>Learning to Rewrite for Non-Autoregressive Neural Machine Translation</i>	
Xinwei Geng, Xiaocheng Feng and Bing Qin	3297
<i>SHAPE: Shifted Absolute Position Embedding for Transformers</i>	
Shun Kiyono, Sosuke Kobayashi, Jun Suzuki and Kentaro Inui	3309
<i>Self-Supervised Quality Estimation for Machine Translation</i>	
Yuanhang Zheng, Zhixing Tan, Meng Zhang, Mieradilijiang Maimaiti, Huanbo Luan, Maosong Sun, Qun Liu and Yang Liu	3322
<i>Generalised Unsupervised Domain Adaptation of Neural Machine Translation with Cross-Lingual Data Selection</i>	
Thuy-Trang Vu, Xuanli He, Dinh Phung and Gholamreza Haffari	3335
<i>STANKER: Stacking Network based on Level-grained Attention-masked BERT for Rumor Detection on Social Media</i>	
Dongning Rao, Xin Miao, Zhihua Jiang and Ran Li	3347

<i>ActiveEA: Active Learning for Neural Entity Alignment</i>	
Bing Liu, Harrison Scells, Guido Zuccon, Wen Hua and Genghong Zhao	3364
<i>Cost-effective End-to-end Information Extraction for Semi-structured Document Images</i>	
Wonseok Hwang, Hyunji Lee, Jinyeong Yim, Geewook Kim and Minjoon Seo	3375
<i>Improving Math Word Problems with Pre-trained Knowledge and Hierarchical Reasoning</i>	
Weijiang Yu, Yingpeng Wen, Fudan Zheng and Nong Xiao	3384
<i>GraphMR: Graph Neural Network for Mathematical Reasoning</i>	
Weijie Feng, Binbin Liu, Dongpeng Xu, Qilong Zheng and Yun Xu	3395
<i>What Changes Can Large-scale Language Models Bring? Intensive Study on HyperCLOVA: Billions-scale Korean Generative Pretrained Transformers</i>	
Boseop Kim, HyoungSeok Kim, Sang-Woo Lee, Gichang Lee, Donghyun Kwak, Jeon Dong Hyeon, Sunghyun Park, Sungju Kim, Seonhoon Kim, Dongpil Seo, Heungsub Lee, Minyoung Jeong, Sungjae Lee, Minsub Kim, Suk Hyun Ko, Seokhun Kim, Taeyong Park, Jinuk Kim, Soyoung Kang, Na-Hyeon Ryu, Kang Min Yoo, Minsuk Chang, Soobin Suh, Sookyo In, Jinseong Park, Kyungduk Kim, Hiun Kim, Jisu Jeong, Yong Goo Yeo, Donghoon Ham, Dongju Park, Min Young Lee, Jaewook Kang, Inho Kang, Jung-Woo Ha, Woomyoung Park and Nako Sung	3405
<i>APIRecX: Cross-Library API Recommendation via Pre-Trained Language Model</i>	
Yuning Kang, Zan Wang, Hongyu Zhang, Junjie Chen and Hanmo You	3425
<i>GMH: A General Multi-hop Reasoning Model for KG Completion</i>	
Yao Zhang, Hongru Liang, Adam Jatowt, Wenqiang Lei, Xin Wei, Ning Jiang and Zhenglu Yang	3437
<i>BPM_MT: Enhanced Backchannel Prediction Model using Multi-Task Learning</i>	
Jin Yea Jang, San Kim, Minyoung Jung, Saim Shin and Gahgene Gweon	3447
<i>Graphine: A Dataset for Graph-aware Terminology Definition Generation</i>	
Zequn Liu, Shukai Wang, Yiyang Gu, Ruiyi Zhang, Ming Zhang and Sheng Wang	3453
<i>Leveraging Order-Free Tag Relations for Context-Aware Recommendation</i>	
Junmo Kang, Jeonghwan Kim, Suwon Shin and Sung-Hyon Myaeng	3464
<i>End-to-End Conversational Search for Online Shopping with Utterance Transfer</i>	
Liqiang Xiao, Jun Ma, Xin Luna Dong, Pascual Martínez-Gómez, Nasser Zalmout, Wei Chen, Tong Zhao, Hao He and Yaohui Jin	3477
<i>Self-Supervised Curriculum Learning for Spelling Error Correction</i>	
Zifa Gan, Hongfei Xu and Hongying Zan	3487
<i>Fix-Filter-Fix: Intuitively Connect Any Models for Effective Bug Fixing</i>	
Haiwen Hong, Jingfeng Zhang, Yin Zhang, Yao Wan and Yulei Sui	3495
<i>Neuro-Symbolic Reinforcement Learning with First-Order Logic</i>	
Daiki Kimura, Masaki Ono, Subhajit Chaudhury, Ryosuke Kohita, Akifumi Wachi, Don Joven Agravante, Michiaki Tsubori, Asim Munawar and Alexander Gray	3505
<i>Biomedical Concept Normalization by Leveraging Hypernyms</i>	
Cheng Yan, Yuanzhe Zhang, Kang Liu, Jun Zhao, Yafei Shi and Shengping Liu	3512

<i>Leveraging Capsule Routing to Associate Knowledge with Medical Literature Hierarchically</i>	
Xin Liu, Qingcai Chen, Junying Chen, Wenxiu Zhou, Tingyu Liu, Xinlan Yang and Weihua Peng	
	3518
<i>Label-Enhanced Hierarchical Contextualized Representation for Sequential Metaphor Identification</i>	
Shuqun Li, Liang Yang, Weidong He, Shiqi Zhang, Jingjie Zeng and Hongfei LIN	3533
<i>SpellBERT: A Lightweight Pretrained Model for Chinese Spelling Check</i>	
Tuo Ji, Hang Yan and Xipeng Qiu	3544
<i>Automated Generation of Accurate & Fluent Medical X-ray Reports</i>	
Hoang Nguyen, Dong Nie, Taivanbat Badamdorj, Yujie Liu, yingying zhu, Jason Truong and Li Cheng	3552
<i>Enhancing Document Ranking with Task-adaptive Training and Segmented Token Recovery Mechanism</i>	
Xingwu Sun, Yanling Cui, Hongyin Tang, Fuzheng Zhang, Beihong Jin and Shi Wang	3570
<i>Abstract, Rationale, Stance: A Joint Model for Scientific Claim Verification</i>	
Zhiwei Zhang, Jiyi Li, Fumiyo Fukumoto and Yanming Ye	3580
<i>A Fine-Grained Domain Adaption Model for Joint Word Segmentation and POS Tagging</i>	
Peijie Jiang, Dingkun Long, Yueheng Sun, Meishan Zhang, Guangwei Xu and Pengjun Xie	3587
<i>Answering Open-Domain Questions of Varying Reasoning Steps from Text</i>	
Peng Qi, Haejun Lee, TG Sido and Christopher Manning	3599
<i>Adaptive Information Seeking for Open-Domain Question Answering</i>	
Yunchang Zhu, Liang Pang, Yanyan Lan, Huawei Shen and Xueqi Cheng	3615
<i>Mapping probability word problems to executable representations</i>	
Simon Suster, Pieter Fivez, Pietro Totis, Angelika Kimmig, Jesse Davis, Luc de Raedt and Walter Daelemans	3627
<i>Enhancing Multiple-choice Machine Reading Comprehension by Punishing Illogical Interpretations</i>	
Yiming Ju, Yuanzhe Zhang, Zhixing Tian, Kang Liu, Xiaohuan Cao, Wenting Zhao, Jinlong li and Jun Zhao	3641
<i>Large-Scale Relation Learning for Question Answering over Knowledge Bases with Pre-trained Language Models</i>	
Yuanmeng Yan, Rumei Li, Sirui Wang, Hongzhi Zhang, Zan Daoguang, Fuzheng Zhang, Wei Wu and Weiran Xu	3653
<i>Phrase Retrieval Learns Passage Retrieval, Too</i>	
Jinhyuk Lee, Alexander Wettig and Danqi Chen	3661
<i>Neural Natural Logic Inference for Interpretable Question Answering</i>	
Jihao Shi, Xiao Ding, Li Du, Ting Liu and Bing Qin	3673
<i>Smoothing Dialogue States for Open Conversational Machine Reading</i>	
Zhuosheng Zhang, Siru Ouyang, Hai Zhao, Masao Utiyama and Eiichiro Sumita	3685
<i>FinQA: A Dataset of Numerical Reasoning over Financial Data</i>	
Zhiyu Chen, Wenhua Chen, Charese Smiley, Sameena Shah, Iana Borova, Dylan Langdon, Reema Moussa, Matt Beane, Ting-Hao Huang, Bryan Routledge and William Yang Wang	3697

<i>FiD-Ex: Improving Sequence-to-Sequence Models for Extractive Rationale Generation</i>	
Kushal Lakhota, Bhargavi Paranjape, Asish Ghoshal, Scott Yih, Yashar Mehdad and Srini Iyer	
3712	
<i>RockNER: A Simple Method to Create Adversarial Examples for Evaluating the Robustness of Named Entity Recognition Models</i>	
Bill Yuchen Lin, Wenyang Gao, Jun Yan, Ryan Moreno and Xiang Ren	3728
<i>Diagnosing the First-Order Logical Reasoning Ability Through LogicNLI</i>	
Jidong Tian, Yitian Li, Wenqing Chen, Liqiang Xiao, Hao He and Yaohui Jin	3738
<i>Constructing a Psychometric Testbed for Fair Natural Language Processing</i>	
Ahmed Abbasi, David Dobolyi, John P. Lalor, Richard G. Netemeyer, Kendall Smith and Yi Yang	
3748	
<i>COUGH: A Challenge Dataset and Models for COVID-19 FAQ Retrieval</i>	
Xinliang Frederick Zhang, Heming Sun, Xiang Yue, Simon Lin and Huan Sun	3759
<i>Chinese WPLC: A Chinese Dataset for Evaluating Pretrained Language Models on Word Prediction Given Long-Range Context</i>	
Huibin Ge, Chenxi Sun, Deyi Xiong and Qun Liu	3770
<i>WinoLogic: A Zero-Shot Logic-based Diagnostic Dataset for Winograd Schema Challenge</i>	
Weinan He, Canming Huang, Yongmei Liu and Xiaodan Zhu	3779
<i>Pseudo Zero Pronoun Resolution Improves Zero Anaphora Resolution</i>	
Ryuto Konno, Shun Kiyono, Yuichiroh Matsubayashi, Hiroki Ouchi and Kentaro Inui	3790
<i>Aligning Cross-lingual Sentence Representations with Dual Momentum Contrast</i>	
Liang Wang, Wei Zhao and Jingming Liu	3807
<i>Total Recall: a Customized Continual Learning Method for Neural Semantic Parsers</i>	
Zhuang Li, Lizhen Qu and Gholamreza Haffari	3816
<i>Exophoric Pronoun Resolution in Dialogues with Topic Regularization</i>	
Xintong Yu, Hongming Zhang, Yangqiu Song, Changshui Zhang, Kun Xu and Dong Yu	3832
<i>Context-Aware Interaction Network for Question Matching</i>	
Zhe Hu, Zuohui Fu, Yu Yin and Gerard de Melo	3846
<i>TEMP: Taxonomy Expansion with Dynamic Margin Loss through Taxonomy-Paths</i>	
Zichen Liu, Hongyuan Xu, Yanlong Wen, Ning Jiang, HaiYing Wu and Xiaojie Yuan	3854
<i>A Graph-Based Neural Model for End-to-End Frame Semantic Parsing</i>	
ZhiChao Lin, Yueheng Sun and Meishan Zhang	3864
<i>Virtual Data Augmentation: A Robust and General Framework for Fine-tuning Pre-trained Models</i>	
Kun Zhou, Wayne Xin Zhao, Sirui Wang, Fuzheng Zhang, Wei Wu and Ji-Rong Wen	3875
<i>CATE: A Contrastive Pre-trained Model for Metaphor Detection with Semi-supervised Learning</i>	
Zhenxi Lin, Qianli Ma, Jiangyue Yan and JIEYU CHEN	3888
<i>To be Closer: Learning to Link up Aspects with Opinions</i>	
Yuxiang Zhou, Lejian Liao, Yang Gao, Zhanming Jie and Wei Lu	3899

<i>Seeking Common but Distinguishing Difference, A Joint Aspect-based Sentiment Analysis Model</i> Hongjiang Jing, Zuchao Li, Hai Zhao and Shu Jiang	3910
<i>Argument Pair Extraction with Mutual Guidance and Inter-sentence Relation Graph</i> Jianzhu Bao, Bin Liang, Jingyi Sun, Yice Zhang, Min Yang and Ruifeng Xu	3923
<i>Emotion Inference in Multi-Turn Conversations with Addressee-Aware Module and Ensemble Strategy</i> Dayu Li, Xiaodan Zhu, Yang Li, Suge Wang, Deyu Li, Jian Liao and Jianxing Zheng	3935
<i>Improving Federated Learning for Aspect-based Sentiment Analysis via Topic Memories</i> Han Qin, Guimin Chen, Yuanhe Tian and Yan Song	3942
<i>Comparative Opinion Quintuple Extraction from Product Reviews</i> Ziheng Liu, Rui Xia and Jianfei Yu	3955
<i>CTAL: Pre-training Cross-modal Transformer for Audio-and-Language Representations</i> Hang Li, Wenbiao Ding, Yu Kang, Tianqiao Liu, Zhongqin Wu and Zitao Liu	3966
<i>Relation-aware Video Reading Comprehension for Temporal Language Grounding</i> Jialin Gao, Xin Sun, Mengmeng Xu, Xi Zhou and Bernard Ghanem	3978
<i>Mutual-Learning Improves End-to-End Speech Translation</i> Jiawei Zhao, Wei Luo, Boxing Chen and Andrew Gilman	3989
<i>Vision Guided Generative Pre-trained Language Models for Multimodal Abstractive Summarization</i> Tiezheng Yu, Wenliang Dai, Zihan Liu and Pascale Fung	3995
<i>Natural Language Video Localization with Learnable Moment Proposals</i> Shaoning Xiao, Long Chen, Jian Shao, Yueting Zhuang and Jun Xiao	4008
<i>Language-Aligned Waypoint (LAW) Supervision for Vision-and-Language Navigation in Continuous Environments</i> Sonia Raychaudhuri, Saim Wani, Shivansh Patel, Unnat Jain and Angel Chang	4018
<i>How to leverage the multimodal EHR data for better medical prediction?</i> Bo Yang and Lijun Wu	4029
<i>Considering Nested Tree Structure in Sentence Extractive Summarization with Pre-trained Transformer</i> Jingun Kwon, Naoki Kobayashi, Hidetaka Kamigaito and Manabu Okumura	4039
<i>Frame Semantic-Enhanced Sentence Modeling for Sentence-level Extractive Text Summarization</i> Yong Guan, Shaoru Guo, Ru Li, Xiaoli Li and Hongye Tan	4045
<i>CAST: Enhancing Code Summarization with Hierarchical Splitting and Reconstruction of Abstract Syntax Trees</i> Ensheng Shi, Yanlin Wang, Lun Du, Hongyu Zhang, Shi Han, Dongmei Zhang and Hongbin Sun	4053
<i>SgSum: Transforming Multi-document Summarization into Sub-graph Selection</i> Moye Chen, Wei Li, Jiachen Liu, Xinyan Xiao, Hua Wu and Haifeng Wang	4063
<i>Event Graph based Sentence Fusion</i> Ruifeng Yuan, zili Wang and Wenjie Li	4075

<i>Transformer-based Lexically Constrained Headline Generation</i>	
Kosuke Yamada, Yuta Hitomi, Hideaki Tamori, Ryohei Sasano, Naoaki Okazaki, Kentaro Inui and Koichi Takeda	4085
<i>Learn to Copy from the Copying History: Correlational Copy Network for Abstractive Summarization</i>	
Haoran Li, Song Xu, Peng Yuan, Yujia Wang, Youzheng Wu, Xiaodong He and Bowen Zhou	4091
<i>Gradient-Based Adversarial Factual Consistency Evaluation for Abstractive Summarization</i>	
Zhiyuan Zeng, Jiaze Chen, Weiran Xu and Lei Li	4102
<i>Word Reordering for Zero-shot Cross-lingual Structured Prediction</i>	
Tao Ji, Yong Jiang, Tao Wang, Zhongqiang Huang, Fei Huang, Yuanbin Wu and Xiaoling Wang	4109
<i>A Unified Encoding of Structures in Transition Systems</i>	
Tao Ji, Yong Jiang, Tao Wang, Zhongqiang Huang, Fei Huang, Yuanbin Wu and Xiaoling Wang	4121
<i>Improving Unsupervised Question Answering via Summarization-Informed Question Generation</i>	
Chenyang Lyu, Lifeng Shang, Yvette Graham, Jennifer Foster, Xin Jiang and Qun Liu	4134
<i>TransferNet: An Effective and Transparent Framework for Multi-hop Question Answering over Relation Graph</i>	
Jiaxin Shi, Shulin Cao, Lei Hou, Juanzi Li and Hanwang Zhang	4149
<i>Topic Transferable Table Question Answering</i>	
Saneem Chemmengath, Vishwajeet Kumar, Samarth Bharadwaj, Jaydeep Sen, Mustafa Canim, Soumen Chakrabarti, Alfio Gliozzo and Karthik Sankaranarayanan	4159
<i>WebSRC: A Dataset for Web-Based Structural Reading Comprehension</i>	
Xingyu Chen, Zihan Zhao, Lu Chen, JiaBao Ji, Danyang Zhang, Ao Luo, Yuxuan Xiong and Kai Yu	4173
<i>Cryptonite: A Cryptic Crossword Benchmark for Extreme Ambiguity in Language</i>	
Avia Efrat, Uri Shaham, Dan Kilman and Omer Levy	4186
<i>End-to-End Entity Resolution and Question Answering Using Differentiable Knowledge Graphs</i>	
Amir Saffari, Armin Oliya, Priyanka Sen and Tom Ayoola	4193
<i>Improving Query Graph Generation for Complex Question Answering over Knowledge Base</i>	
Kechen Qin, Cheng Li, Virgil Pavlu and Javed Aslam	4201
<i>DiscoDVT: Generating Long Text with Discourse-Aware Discrete Variational Transformer</i>	
Haozhe Ji and Minlie Huang	4208
<i>Mathematical Word Problem Generation from Commonsense Knowledge Graph and Equations</i>	
Tianqiao Liu, Qiang Fang, Wenbiao Ding, Hang Li, Zhongqin Wu and Zitao Liu	4225
<i>Generic resources are what you need: Style transfer tasks without task-specific parallel training data</i>	
Huiyuan Lai, Antonio Toral and Malvina Nissim	4241
<i>Revisiting Pivot-Based Paraphrase Generation: Language Is Not the Only Optional Pivot</i>	
Yitao Cai, Yue Cao and Xiaojun Wan	4255
<i>Structural Adapters in Pretrained Language Models for AMR-to-Text Generation</i>	
Leonardo F. R. Ribeiro, Yue Zhang and Iryna Gurevych	4269

<i>Data-to-text Generation by Splicing Together Nearest Neighbors</i>	
Sam Wiseman, Arturs Backurs and Karl Stratos	4283
<i>Contextualize Knowledge Bases with Transformer for End-to-end Task-Oriented Dialogue Systems</i>	
Yanjie Gou, Yinjie Lei, Lingqiao Liu, Yong Dai and Chunxu Shen	4300
<i>Efficient Dialogue Complementary Policy Learning via Deep Q-network Policy and Episodic Memory Policy</i>	
Yangyang Zhao, Zhenyu Wang, Changxi Zhu and Shihan Wang	4311
<i>CRFR: Improving Conversational Recommender Systems via Flexible Fragments Reasoning on Knowledge Graphs</i>	
Jinfeng Zhou, Bo Wang, Ruifang He and Yuexian Hou	4324
<i>DuRecDial 2.0: A Bilingual Parallel Corpus for Conversational Recommendation</i>	
Zeming Liu, Haifeng Wang, Zheng-Yu Niu, Hua Wu and Wanxiang Che	4335
<i>End-to-End Learning of Flowchart Grounded Task-Oriented Dialogs</i>	
Dinesh Raghu, Shantanu Agarwal, Sachindra Joshi and	4348
<i>Dimensional Emotion Detection from Categorical Emotion</i>	
Sungjoon Park, Jiseon Kim, Seonghyeon Ye, Jaeyeol Jeon, Hee Young Park and Alice Oh ...	4367
<i>Not All Negatives are Equal: Label-Aware Contrastive Loss for Fine-grained Text Classification</i>	
Varsha Suresh and Desmond Ong	4381
<i>Joint Multi-modal Aspect-Sentiment Analysis with Auxiliary Cross-modal Relation Detection</i>	
Xincheng Ju, Dong Zhang, Rong Xiao, Junhui Li, Shoushan Li, Min Zhang and Guodong Zhou	4395
<i>Solving Aspect Category Sentiment Analysis as a Text Generation Task</i>	
Jian Liu, Zhiyang Teng, Leyang Cui, Hanmeng Liu and Yue Zhang	4406
<i>Semantics-Preserved Data Augmentation for Aspect-Based Sentiment Analysis</i>	
Ting-Wei Hsu, Chung-Chi Chen, Hen-Hsen Huang and Hsin-Hsi Chen	4417
<i>The Effect of Round-Trip Translation on Fairness in Sentiment Analysis</i>	
Jonathan Christiansen, Mathias Gammelgaard and Anders Sjøgaard	4423
<i>CHoRaL: Collecting Humor Reaction Labels from Millions of Social Media Users</i>	
Zixiaofan Yang, Shayan Hooshmand and Julia Hirschberg	4429
<i>CSDS: A Fine-Grained Chinese Dataset for Customer Service Dialogue Summarization</i>	
Haitao Lin, Liqun Ma, Junnan Zhu, Lu Xiang, Yu Zhou, Jiajun Zhang and Chengqing Zong .	4436
<i>CodRED: A Cross-Document Relation Extraction Dataset for Acquiring Knowledge in the Wild</i>	
Yuan Yao, Jiaju Du, Yankai Lin, Peng Li, Zhiyuan Liu, Jie Zhou and Maosong Sun	4452
<i>Building and Evaluating Open-Domain Dialogue Corpora with Clarifying Questions</i>	
Mohammad Aliannejadi, Julia Kiseleva, Aleksandr Chuklin, Jeff Dalton and Mikhail Burtsev	4473
<i>We Need to Talk About train-dev-test Splits</i>	
Rob van der Goot	4485

<i>PhoMT: A High-Quality and Large-Scale Benchmark Dataset for Vietnamese-English Machine Translation</i>	
Long Doan, Linh The Nguyen, Nguyen Luong Tran, Thai Hoang and Dat Quoc Nguyen	4495
<i>Lying Through One’s Teeth: A Study on Verbal Leakage Cues</i>	
Min-Hsuan Yeh and Lun-Wei Ku	4504
<i>Multi-granularity Textual Adversarial Attack with Behavior Cloning</i>	
Yangyi Chen, Jin Su and Wei Wei	4511
<i>All Bark and No Bite: Rogue Dimensions in Transformer Language Models Obscure Representational Quality</i>	
William Timkey and Marten van Schijndel	4527
<i>Incorporating Residual and Normalization Layers into Analysis of Masked Language Models</i>	
Goro Kobayashi, Tatsuki Kuribayashi, Sho Yokoi and Kentaro Inui	4547
<i>Mind the Style of Text! Adversarial and Backdoor Attacks Based on Text Style Transfer</i>	
Fanchao Qi, Yangyi Chen, Xurui Zhang, Mukai Li, Zhiyuan Liu and Maosong Sun	4569
<i>Sociolectal Analysis of Pretrained Language Models</i>	
Sheng Zhang, Xin Zhang, Weiming Zhang and Anders Søgaard	4581
<i>Examining Cross-lingual Contextual Embeddings with Orthogonal Structural Probes</i>	
Tomasz Limisiewicz and David Mareček	4589
<i>Are Transformers a Modern Version of ELIZA? Observations on French Object Verb Agreement</i>	
Bingzhi Li, Guillaume Wisniewski and Benoit Crabbé	4599
<i>Fine-grained Entity Typing via Label Reasoning</i>	
Qing Liu, Hongyu Lin, Xinyan Xiao, Xianpei Han, Le Sun and Hua Wu	4611
<i>Enhanced Language Representation with Label Knowledge for Span Extraction</i>	
Pan Yang, Xin Cong, Zhenyu Sun and Xingwu Liu	4623
<i>PRIDE: Predicting Relationships in Conversations</i>	
Anna Tiginova, Paramita Mirza, Andrew Yates and Gerhard Weikum	4636
<i>Extracting Fine-Grained Knowledge Graphs of Scientific Claims: Dataset and Transformer-Based Results</i>	
Ian Magnusson and Scott Friedman	4651
<i>Sequential Cross-Document Coreference Resolution</i>	
Emily Allaway, Shuai Wang and Miguel Ballesteros	4659
<i>Mixture-of-Partitions: Infusing Large Biomedical Knowledge Graphs into BERT</i>	
Zaiqiao Meng, Fangyu Liu, Thomas Clark, Ehsan Shareghi and Nigel Collier	4672
<i>Filling the Gaps in Ancient Akkadian Texts: A Masked Language Modelling Approach</i>	
Koren Lazar, Benny Saret, Asaf Yehudai, Wayne Horowitz, Nathan Wasserman and Gabriel Stanovsky	4682
<i>AVocaDo: Strategy for Adapting Vocabulary to Downstream Domain</i>	
Jimin Hong, TaeHee Kim, Hyesu Lim and Jaegul Choo	4692

<i>Can We Improve Model Robustness through Secondary Attribute Counterfactuals?</i>	
Ananth Balashankar, Xuezhong Wang, Ben Packer, Nithum Thain, Ed Chi and Alex Beutel	4701
<i>Long-Range Modeling of Source Code Files with eWASH: Extended Window Access by Syntax Hierarchy</i>	
Colin Clement, Shuai Lu, Xiaoyu Liu, Michele Tufano, Dawn Drain, Nan Duan, Neel Sundaresan and Alexey Svyatkovskiy	4713
<i>Can Language Models be Biomedical Knowledge Bases?</i>	
Mujeen Sung, Jinhyuk Lee, Sean Yi, Minji Jeon, Sungdong Kim and Jaewoo Kang	4723
<i>LayoutReader: Pre-training of Text and Layout for Reading Order Detection</i>	
Zilong Wang, Yiheng Xu, Lei Cui, Jingbo Shang and Furu Wei	4735
<i>Region under Discussion for visual dialog</i>	
Mauricio Mazuecos, Franco M. Luque, Jorge Sánchez, Hernán Maina, Thomas Vadora and Luciana Benotti	4745
<i>Learning grounded word meaning representations on similarity graphs</i>	
Mariella Dimiccoli, Herwig Wendt and Pau Batlle Franch	4760
<i>WhyAct: Identifying Action Reasons in Lifestyle Vlogs</i>	
Oana Ignat, Santiago Castro, Hanwen Miao, Weiji Li and Rada Mihalcea	4770
<i>Genre as Weak Supervision for Cross-lingual Dependency Parsing</i>	
Max Müller-Eberstein, Rob van der Goot and Barbara Plank	4786
<i>On the Relation between Syntactic Divergence and Zero-Shot Performance</i>	
Ofir Arviv, Dmitry Nikolaev, Taelin Karidi and Omri Abend	4803
<i>Improved Latent Tree Induction with Distant Supervision via Span Constraints</i>	
Zhiyang Xu, Andrew Drozdov, Jay Yoon Lee, Tim O’Gorman, Subendhu Rongali, Dylan Finkbeiner, Shilpa Suresh, Mohit Iyyer and Andrew McCallum	4818
<i>Aligning Multidimensional Worldviews and Discovering Ideological Differences</i>	
Jeremiah Milbauer, Adarsh Mathew and James Evans	4832
<i>Just Say No: Analyzing the Stance of Neural Dialogue Generation in Offensive Contexts</i>	
Ashutosh Baheti, Maarten Sap, Alan Ritter and Mark Riedl	4846
<i>Multi-Modal Open-Domain Dialogue</i>	
Kurt Shuster, Eric Michael Smith, Da Ju and Jason Weston	4863
<i>A Label-Aware BERT Attention Network for Zero-Shot Multi-Intent Detection in Spoken Language Understanding</i>	
Ting-Wei Wu, Ruolin Su and Biing Juang	4884
<i>Zero-Shot Dialogue Disentanglement by Self-Supervised Entangled Response Selection</i>	
Ta-Chung Chi and alexander rudnicky	4897
<i>SIMMC 2.0: A Task-oriented Dialog Dataset for Immersive Multimodal Conversations</i>	
Satwik Kottur, Seungwhan Moon, Alborz Geramifard and Babak Damavandi	4903
<i>RAST: Domain-Robust Dialogue Rewriting as Sequence Tagging</i>	
Jie Hao, Linfeng Song, Liwei Wang, Kun Xu, Zhaopeng Tu and Dong Yu	4913

<i>MRF-Chat: Improving Dialogue with Markov Random Fields</i>	
Ishaan Grover, Matthew Huggins, Cynthia Breazeal and Hae Won Park	4925
<i>Dialogue State Tracking with a Language Model using Schema-Driven Prompting</i>	
Chia-Hsuan Lee, Hao Cheng and Mari Ostendorf	4937
<i>Signed Coreference Resolution</i>	
Kayo Yin, Kenneth DeHaan and Malihe Alikhani	4950
<i>Consistent Accelerated Inference via Confident Adaptive Transformers</i>	
Tal Schuster, Adam Fisch, Tommi Jaakkola and Regina Barzilay	4962
<i>Improving and Simplifying Pattern Exploiting Training</i>	
Derek Tam, Rakesh R. Menon, Mohit Bansal, Shashank Srivastava and Colin Raffel	4980
<i>Unsupervised Data Augmentation with Naive Augmentation and without Unlabeled Data</i>	
David Lowell, Brian Howard, Zachary C. Lipton and Byron Wallace	4992
<i>Pre-train or Annotate? Domain Adaptation with a Constrained Budget</i>	
Fan Bai, Alan Ritter and Wei Xu	5002
<i>Lawyers are Dishonest? Quantifying Representational Harms in Commonsense Knowledge Resources</i>	
Ninareh Mehrabi, Pei Zhou, Fred Morstatter, Jay Pujara, Xiang Ren and Aram Galstyan	5016
<i>OSCaR: Orthogonal Subspace Correction and Rectification of Biases in Word Embeddings</i>	
Sunipa Dev, Tao Li, Jeff M Phillips and Vivek Srikumar	5034
<i>Sentence-Permuted Paragraph Generation</i>	
Wenhao Yu, Chenguang Zhu, Tong Zhao, Zhichun Guo and Meng Jiang	5051
<i>Extract, Denoise and Enforce: Evaluating and Improving Concept Preservation for Text-to-Text Generation</i>	
Yuning Mao, Wenchang Ma, Deren Lei, Jiawei Han and Xiang Ren	5063
<i>Paraphrase Generation: A Survey of the State of the Art</i>	
Jianing Zhou and Suma Bhat	5075
<i>Exposure Bias versus Self-Recovery: Are Distortions Really Incremental for Autoregressive Text Generation?</i>	
Tianxing He, Jingzhao Zhang, Zhiming Zhou and James Glass	5087
<i>Generating Self-Contained and Summary-Centric Question Answer Pairs via Differentiable Reward Imitation Learning</i>	
Li Zhou, Kevin Small, Yong Zhang and Sandeep Atluri	5103
<i>Unsupervised Paraphrasing with Pretrained Language Models</i>	
Tong Niu, Semih Yavuz, Yingbo Zhou, Nitish Shirish Keskar, Huan Wang and Caiming Xiong	5136
<i>Profanity-Avoiding Training Framework for Seq2seq Models with Certified Robustness</i>	
Hengtong Zhang, Tianhang Zheng, Yaliang Li, Jing Gao, Lu Su and Bo Li	5151
<i>Journalistic Guidelines Aware News Image Captioning</i>	
Xuewen Yang, Svebor Karaman, Joel Tetreault and Alejandro Jaimes	5162
<i>AESOP: Paraphrase Generation with Adaptive Syntactic Control</i>	
Jiao Sun, Xuezhe Ma and Nanyun Peng	5176

<i>Refocusing on Relevance: Personalization in NLG</i>	
Shiran Dudy, Steven Bedrick and Bonnie Webber	5190
<i>The Future is not One-dimensional: Complex Event Schema Induction by Graph Modeling for Event Prediction</i>	
Manling Li, Sha Li, Zhenhailong Wang, Lifu Huang, Kyunghyun Cho, Heng Ji, Jiawei Han and Clare Voss	5203
<i>Learning Constraints and Descriptive Segmentation for Subevent Detection</i>	
Haoyu Wang, Hongming Zhang, Muhao Chen and Dan Roth	5216
<i>ChemNER: Fine-Grained Chemistry Named Entity Recognition with Ontology-Guided Distant Supervision</i>	
Xuan Wang, Vivian Hu, Xiangchen Song, Shweta Garg, Jinfeng Xiao and Jiawei Han	5227
<i>Moving on from OntoNotes: Coreference Resolution Model Transfer</i>	
Patrick Xia and Benjamin Van Durme	5241
<i>Document-level Entity-based Extraction as Template Generation</i>	
Kung-Hsiang Huang, Sam Tang and Nanyun Peng	5257
<i>Learning Prototype Representations Across Few-Shot Tasks for Event Detection</i>	
Viet Lai, Franck Dernoncourt and Thien Huu Nguyen	5270
<i>Lifelong Event Detection with Knowledge Transfer</i>	
Pengfei Yu, Heng Ji and Prem Natarajan	5278
<i>Modular Self-Supervision for Document-Level Relation Extraction</i>	
Sheng Zhang, Cliff Wong, Naoto Usuyama, Sarthak Jain, Tristan Naumann and Hoifung Poon	5291
<i>Unsupervised Paraphrasing Consistency Training for Low Resource Named Entity Recognition</i>	
Rui Wang and Ricardo Henao	5303
<i>Fine-grained Entity Typing without Knowledge Base</i>	
Jing Qian, Yibin Liu, Lemao Liu, Yangming Li, Haiyun Jiang, Haisong Zhang and Shuming Shi	5309
<i>Adversarial Attack against Cross-lingual Knowledge Graph Alignment</i>	
Zeru Zhang, Zijie Zhang, Yang Zhou, Lingfei Wu, Sixing Wu, Xiaoying Han, Dejing Dou, Tianshi Che and Da Yan	5320
<i>Towards Realistic Few-Shot Relation Extraction</i>	
Sam Brody, Sichao Wu and Adrian Benton	5338
<i>Data Augmentation for Cross-Domain Named Entity Recognition</i>	
Shuguang Chen, Gustavo Aguilar, Leonardo Neves and Thamar Solorio	5346
<i>Incorporating medical knowledge in BERT for clinical relation extraction</i>	
Arpita Roy and Shimei Pan	5357
<i>ECONET: Effective Continual Pretraining of Language Models for Event Temporal Reasoning</i>	
Rujun Han, Xiang Ren and Nanyun Peng	5367
<i>Learning from Noisy Labels for Entity-Centric Information Extraction</i>	
Wenxuan Zhou and Muhao Chen	5381

<i>Extracting Material Property Measurement Data from Scientific Articles</i> Gihan Panapitiya, Fred Parks, Jonathan Sepulveda and Emily Saldanha	5393
<i>Modeling Document-Level Context for Event Detection via Important Context Selection</i> Amir Pouran Ben Veysch, Minh Van Nguyen, Nghia Ngo Trung, Bonan Min and Thien Huu Nguyen 5403	
<i>Crosslingual Transfer Learning for Relation and Event Extraction via Word Category and Class Alignments</i> Minh Van Nguyen, Tuan Ngo Nguyen, Bonan Min and Thien Huu Nguyen	5414
<i>Corpus-based Open-Domain Event Type Induction</i> Jiaming Shen, Yunyi Zhang, Heng Ji and Jiawei Han	5427
<i>PDALN: Progressive Domain Adaptation over a Pre-trained Model for Low-Resource Cross-Domain Named Entity Recognition</i> Tao Zhang, Congying Xia, Philip S. Yu, Zhiwei Liu and Shu Zhao	5441
<i>Multi-Vector Attention Models for Deep Re-ranking</i> Giulio Zhou and Jacob Devlin	5452
<i>Toward Deconfounding the Effect of Entity Demographics for Question Answering Accuracy</i> Maharshi Gor, Kellie Webster and Jordan Boyd-Graber	5457
<i>Exploring Strategies for Generalizable Commonsense Reasoning with Pre-trained Models</i> Kaixin Ma, Filip Ilievski, Jonathan Francis, Satoru Ozaki, Eric Nyberg and Alessandro Oltramari 5474	
<i>Transformer Feed-Forward Layers Are Key-Value Memories</i> Mor Geva, Roei Schuster, Jonathan Berant and Omer Levy	5484
<i>Connecting Attributions and QA Model Behavior on Realistic Counterfactuals</i> Xi Ye, Rohan Nair and Greg Durrett	5496
<i>How Do Neural Sequence Models Generalize? Local and Global Cues for Out-of-Distribution Prediction</i> D. Anthony Bau and Jacob Andreas	5513
<i>Comparing Text Representations: A Theory-Driven Approach</i> Gregory Yauney and David Mimno	5527
<i>Human Rationales as Attribution Priors for Explainable Stance Detection</i> Sahil Jayaram and Emily Allaway	5540
<i>The Stem Cell Hypothesis: Dilemma behind Multi-Task Learning with Transformer Encoders</i> Han He and Jinho D. Choi	5555
<i>Text Counterfactuals via Latent Optimization and Shapley-Guided Search</i> Xiaoli Fern and Quintin Pope	5578
<i>“Average” Approximates “First Principal Component”? An Empirical Analysis on Representations from Neural Language Models</i> Zihan Wang, Chengyu Dong and Jingbo Shang	5594
<i>Controlled Evaluation of Grammatical Knowledge in Mandarin Chinese Language Models</i> Yiwen Wang, Jennifer Hu, Roger Levy and Peng Qian	5604

<i>GradTS: A Gradient-Based Automatic Auxiliary Task Selection Method Based on Transformer Networks</i> Weicheng Ma, Renze Lou, Kai Zhang, Lili Wang and Soroush Vosoughi	5621
<i>NegatER: Unsupervised Discovery of Negatives in Commonsense Knowledge Bases</i> Tara Safavi, Jing Zhu and Danai Koutra	5633
<i>Instance-adaptive training with noise-robust losses against noisy labels</i> Lifeng Jin, Linfeng Song, Kun Xu and Dong Yu	5647
<i>Distributionally Robust Multilingual Machine Translation</i> Chunting Zhou, Daniel Levy, Xian Li, Marjan Ghazvininejad and Graham Neubig	5664
<i>Model Selection for Cross-lingual Transfer</i> Yang Chen and Alan Ritter	5675
<i>Continual Few-Shot Learning for Text Classification</i> Ramakanth Pasunuru, Veselin Stoyanov and Mohit Bansal	5688
<i>Efficient Nearest Neighbor Language Models</i> Junxian He, Graham Neubig and Taylor Berg-Kirkpatrick	5703
<i>STraTA: Self-Training with Task Augmentation for Better Few-shot Learning</i> Tu Vu, Minh-Thang Luong, Quoc Le, Grady Simon and Mohit Iyyer	5715
<i>TADPOLE: Task ADapted Pre-Training via AnOmaLy DETection</i> Vivek Madan, Ashish Khetan and Zohar Karnin	5732
<i>Gradient-based Adversarial Attacks against Text Transformers</i> Chuan Guo, Alexandre Sablayrolles, Hervé Jégou and Douwe Kiela	5747
<i>Do Transformer Modifications Transfer Across Implementations and Applications?</i> Sharan Narang, Hyung Won Chung, Yi Tay, Liam Fedus, Thibault Fevry, Michael Matena, Karishma Malkan, Noah Fiedel, Noam Shazeer, Zhenzhong Lan, Yanqi Zhou, Wei Li, Nan Ding, Jake Marcus, Adam Roberts and Colin Raffel	5758
<i>Paired Examples as Indirect Supervision in Latent Decision Models</i> Nitish Gupta, Sameer Singh, Matt Gardner and Dan Roth	5774
<i>Pairwise Supervised Contrastive Learning of Sentence Representations</i> Dejiao Zhang, Shang-Wen Li, Wei Xiao, Henghui Zhu, Ramesh Nallapati, Andrew O. Arnold and Bing Xiang	5786
<i>Muppet: Massive Multi-task Representations with Pre-Finetuning</i> Armen Aghajanyan, Anchit Gupta, Akshat Shrivastava, Xilun Chen, Luke Zettlemoyer and Sonal Gupta	5799
<i>Diverse Distributions of Self-Supervised Tasks for Meta-Learning in NLP</i> Trapit Bansal, Karthick Prasad Gunasekaran, Tong Wang, Tsendsuren Munkhdalai and Andrew McCallum	5812
<i>A Simple and Effective Method To Eliminate the Self Language Bias in Multilingual Representations</i> Ziyi Yang, Yinfei Yang, Daniel Cer and Eric Darve	5825
<i>A Massively Multilingual Analysis of Cross-linguality in Shared Embedding Space</i> Alexander Jones, William Yang Wang and Kyle Mahowald	5833

<i>Frustratingly Simple but Surprisingly Strong: Using Language-Independent Features for Zero-shot Cross-lingual Semantic Parsing</i>	
Jingfeng Yang, Federico Fancellu, Bonnie Webber and Diyi Yang	5848
<i>Improving Simultaneous Translation by Incorporating Pseudo-References with Fewer Reorderings</i>	
Junkun Chen, Renjie Zheng, Atsuhito Kita, Mingbo Ma and Liang Huang	5857
<i>Classification-based Quality Estimation: Small and Efficient Models for Real-world Applications</i>	
Shuo Sun, Ahmed El-Kishky, Vishrav Chaudhary, James Cross, Lucia Specia and Francisco Guzmán	5865
<i>A Large-Scale Study of Machine Translation in Turkic Languages</i>	
Jamshidbek Mirzakhalov, Anoop Babu, Duygu Ataman, Sherzod Kariev, Francis Tyers, Otabek Abduraufov, Mammad Hajili, Sardana Ivanova, Abror Khaytbaev, Antonio Laverghetta Jr., Bekhzodbek Moydinboev, Esra Onal, Shaxnoza Pulatova, Ahsan Wahab, Orhan Firat and Sriram Chellappan .	5876
<i>Analyzing the Surprising Variability in Word Embedding Stability Across Languages</i>	
Laura Burdick, Jonathan K. Kummerfeld and Rada Mihalcea	5891
<i>Rule-based Morphological Inflection Improves Neural Terminology Translation</i>	
Weijia Xu and Marine Carpuat	5902
<i>Data and Parameter Scaling Laws for Neural Machine Translation</i>	
Mitchell A Gordon, Kevin Duh and Jared Kaplan	5915
<i>Good-Enough Example Extrapolation</i>	
Jason Wei	5923
<i>Learning to Selectively Learn for Weakly-supervised Paraphrase Generation</i>	
Kaize Ding, Dingcheng Li, Alexander Hanbo Li, Xing Fan, Chenlei Guo, Yang Liu and Huan Liu	5930
<i>Effective Convolutional Attention Network for Multi-label Clinical Document Classification</i>	
Yang Liu, Hua Cheng, Russell Klopfer, Matthew R. Gormley and Thomas Schaaf	5941
<i>Contrastive Code Representation Learning</i>	
Paras Jain, Ajay Jain, Tianjun Zhang, Pieter Abbeel, Joseph Gonzalez and Ion Stoica	5954
<i>IGA: An Intent-Guided Authoring Assistant</i>	
Simeng Sun, Wenlong Zhao, Varun Manjunatha, Rajiv Jain, Vlad Morariu, Franck Dernoncourt, Balaji Vasani Srinivasan and Mohit Iyyer	5972
<i>Math Word Problem Generation with Mathematical Consistency and Problem Context Constraints</i>	
Zichao Wang, Andrew Lan and Richard Baraniuk	5986
<i>Navigating the Kaleidoscope of COVID-19 Misinformation Using Deep Learning</i>	
Yuanzhi Chen and Mohammad Hasan	6000
<i>Detecting Health Advice in Medical Research Literature</i>	
Yingya Li, Jun Wang and Bei Yu	6018
<i>A Semantic Feature-Wise Transformation Relation Network for Automatic Short Answer Grading</i>	
Zhaohui Li, Yajur Tomar and Rebecca J. Passonneau	6030
<i>Evaluating Scholarly Impact: Towards Content-Aware Bibliometrics</i>	
Saurav Manchanda and George Karypis	6041

<i>A Scalable Framework for Learning From Implicit User Feedback to Improve Natural Language Understanding in Large-Scale Conversational AI Systems</i>	
Sunghyun Park, Han Li, Ameen Patel, Sidharth Mudgal, Sungjin Lee, Young-Bum Kim, Spyros Matsoukas and Ruhi Sarikaya	6054
<i>Summarize-then-Answer: Generating Concise Explanations for Multi-hop Reading Comprehension</i>	
Naoya Inoue, Harsh Trivedi, Steven Sinha, Niranjana Balasubramanian and Kentaro Inui	6064
<i>FewshotQA: A simple framework for few-shot learning of question answering tasks using pre-trained text-to-text models</i>	
Rakesh Chada and Pradeep Natarajan	6081
<i>Multi-stage Training with Improved Negative Contrast for Neural Passage Retrieval</i>	
Jing Lu, Gustavo Hernandez Abrego, Ji Ma, Jianmo Ni and Yinfei Yang	6091
<i>Perhaps PTLMs Should Go to School – A Task to Assess Open Book and Closed Book QA</i>	
Manuel Ciosici, Joe Cecil, Dong-Ho Lee, Alex Hedges, Marjorie Freedman and Ralph Weischedel	6104
<i>ReasonBERT: Pre-trained to Reason with Distant Supervision</i>	
Xiang Deng, Yu Su, Alyssa Lees, You Wu, Cong Yu and Huan Sun	6112
<i>Single-dataset Experts for Multi-dataset Question Answering</i>	
Dan Friedman, Ben Dodge and Danqi Chen	6128
<i>Simple Entity-Centric Questions Challenge Dense Retrievers</i>	
Christopher Scialolino, Zexuan Zhong, Jinhyuk Lee and Danqi Chen	6138
<i>Mitigating False-Negative Contexts in Multi-document Question Answering with Retrieval Marginalization</i>	
Ansong Ni, Matt Gardner and Pradeep Dasigi	6149
<i>MultiDoc2Dial: Modeling Dialogues Grounded in Multiple Documents</i>	
Song Feng, Siva Sankalp Patel, Hui Wan and Sachindra Joshi	6162
<i>GupShup: Summarizing Open-Domain Code-Switched Conversations</i>	
Laiba Mehnaz, Debanjan Mahata, Rakesh Gosangi, Uma Sushmitha Gunturi, Riya Jain, Gauri Gupta, Amardeep Kumar, Isabelle G. Lee, Anish Acharya and Rajiv Ratn Shah	6177
<i>BiSECT: Learning to Split and Rephrase Sentences with Bitexts</i>	
Joongwon Kim, Mounica Maddela, Reno Kriz, Wei Xu and Chris Callison-Burch	6193
<i>Data Collection vs. Knowledge Graph Completion: What is Needed to Improve Coverage?</i>	
Kenneth Church and Yuchen Bian	6210
<i>Universal Sentence Representation Learning with Conditional Masked Language Model</i>	
Ziyi Yang, Yinfei Yang, Daniel Cer, Jax Law and Eric Darve	6216
<i>On the Benefit of Syntactic Supervision for Cross-lingual Transfer in Semantic Role Labeling</i>	
Zhisong Zhang, Emma Strubell and Eduard Hovy	6229
<i>Implicit Premise Generation with Discourse-aware Commonsense Knowledge Models</i>	
Tuhin Chakrabarty, Aadit Trivedi and Smaranda Muresan	6247
<i>Inducing Transformer’s Compositional Generalization Ability via Auxiliary Sequence Prediction Tasks</i>	
Yichen Jiang and Mohit Bansal	6253

<i>Flexible Generation of Natural Language Deductions</i>	
Kaj Bostrom, Xinyu Zhao, Swarat Chaudhuri and Greg Durrett	6266
<i>Structure-aware Fine-tuning of Sequence-to-sequence Transformers for Transition-based AMR Parsing</i>	
Jiawei Zhou, Tahira Naseem, Ramón Fernandez Astudillo, Young-Suk Lee, Radu Florian and Salim Roukos.....	6279
<i>Think about it! Improving defeasible reasoning by first modeling the question scenario.</i>	
Aman Madaan, Niket Tandon, Dheeraj Rajagopal, Peter Clark, Yiming Yang and Eduard Hovy	6291
<i>Open Aspect Target Sentiment Classification with Natural Language Prompts</i>	
Ronald Seoh, Ian Birle, Mrinal Tak, Haw-Shiuan Chang, Brian Pinette and Alfred Hough....	6311
<i>Does BERT Learn as Humans Perceive? Understanding Linguistic Styles through Lexica</i>	
Shirley Hayati, Dongyeop Kang and Lyle Ungar	6323
<i>Improving Stance Detection with Multi-Dataset Learning and Knowledge Distillation</i>	
Yingjie Li, Chenye Zhao and Cornelia Caragea	6332
<i>Discovering the Unknown Knowns: Turning Implicit Knowledge in the Dataset into Explicit Training Examples for Visual Question Answering</i>	
Jihyung Kil, Cheng Zhang, Dong Xuan and Wei-Lun Chao	6346
<i>Improving Pre-trained Vision-and-Language Embeddings for Phrase Grounding</i>	
Zi-Yi Dou and Nanyun Peng	6362
<i>Sequential Randomized Smoothing for Adversarially Robust Speech Recognition</i>	
Raphael Olivier and Bhiksha Raj	6372
<i>Hitting your MARQ: Multimodal ARgument Quality Assessment in Long Debate Video</i>	
Md Kamrul Hasan, James Spann, Masum Hasan, Md Saiful Islam, Kurtis Haut, Rada Mihalcea and Ehsan Hoque	6387
<i>Mind the Context: The Impact of Contextualization in Neural Module Networks for Grounding Visual Referring Expressions</i>	
Arjun Akula, Spandana Gella, Keze Wang, Song-Chun Zhu and Siva Reddy	6398
<i>Weakly-Supervised Visual-Retriever-Reader for Knowledge-based Question Answering</i>	
Man Luo, Yankai Zeng, Pratyay Banerjee and Chitta Baral	6417
<i>NDH-Full: Learning and Evaluating Navigational Agents on Full-Length Dialogue</i>	
Hyoungun Kim, Jialu Li and Mohit Bansal.....	6432
<i>Timeline Summarization based on Event Graph Compression via Time-Aware Optimal Transport</i>	
Manling Li, Tengfei Ma, Mo Yu, Lingfei Wu, Tian Gao, Heng Ji and Kathleen McKeown ...	6443
<i>StreamHover: Livestream Transcript Summarization and Annotation</i>	
Sangwoo Cho, Franck Dernoncourt, Tim Ganter, Trung Bui, Nedim Lipka, Walter Chang, Hailin Jin, Jonathan Brandt, Hassan Foroosh and Fei Liu	6457
<i>Cross-Register Projection for Headline Part of Speech Tagging</i>	
Adrian Benton, Hanyang Li and Igor Malioutov	6475
<i>Editing Factual Knowledge in Language Models</i>	
Nicola De Cao, Wilker Aziz and Ivan Titov	6491

<i>Sparse Attention with Linear Units</i>	
Biao Zhang, Ivan Titov and Rico Sennrich	6507
<i>Knowledge Base Completion Meets Transfer Learning</i>	
Vid Kocijan and Thomas Lukasiewicz	6521
<i>SPECTRA: Sparse Structured Text Rationalization</i>	
Nuno M. Guerreiro and André F. T. Martins	6534
<i>Towards Zero-Shot Knowledge Distillation for Natural Language Processing</i>	
Ahmad Rashid, Vasileios Lioutas, Abbas Ghaddar and Mehdi Rezagholizadeh	6551
<i>Adversarial Regularization as Stackelberg Game: An Unrolled Optimization Approach</i>	
Simiao Zuo, Chen Liang, Haoming Jiang, Xiaodong Liu, Pengcheng He, Jianfeng Gao, Weizhu Chen and Tuo Zhao	6562
<i>Aspect-Controllable Opinion Summarization</i>	
Reinald Kim Amplayo, Stefanos Angelidis and Mirella Lapata	6578
<i>QuestEval: Summarization Asks for Fact-based Evaluation</i>	
Thomas Scialom, Paul-Alexis Dray, Sylvain Lamprier, Benjamin Piwowarski, Jacopo Staiano, Alex Wang and patrick Gallinari	6594
<i>Simple Conversational Data Augmentation for Semi-supervised Abstractive Dialogue Summarization</i>	
Jiaao Chen and Diyi Yang	6605
<i>Finding a Balanced Degree of Automation for Summary Evaluation</i>	
Shiyue Zhang and Mohit Bansal	6617
<i>CLIFF: Contrastive Learning for Improving Faithfulness and Factuality in Abstractive Summarization</i>	
Shuyang Cao and Lu Wang	6633
<i>Multilingual Unsupervised Neural Machine Translation with Denoising Adapters</i>	
Ahmet Üstün, Alexandre Berard, Laurent Besacier and Matthias Gallé	6650
<i>BERT, mBERT, or BiBERT? A Study on Contextualized Embeddings for Neural Machine Translation</i>	
Haoran Xu, Benjamin Van Durme and Kenton Murray	6663
<i>Controlling Machine Translation for Multiple Attributes with Additive Interventions</i>	
Andrea Schioppa, David Vilar, Artem Sokolov and Katja Filippova	6676
<i>A Generative Framework for Simultaneous Machine Translation</i>	
Yishu Miao, Phil Blunsom and Lucia Specia	6697
<i>It Is Not As Good As You Think! Evaluating Simultaneous Machine Translation on Interpretation Data</i>	
Jinming Zhao, Philip Arthur, Gholamreza Haffari, Trevor Cohn and Ehsan Shareghi	6707
<i>Boosting Cross-Lingual Transfer via Self-Learning with Uncertainty Estimation</i>	
Liyang Xu, Xuchao Zhang, Xujiang Zhao, Haifeng Chen, Feng Chen and Jinho D. Choi	6716
<i>Levenshtein Training for Word-level Quality Estimation</i>	
Shuoyang Ding, Marcin Junczys-Dowmunt, Matt Post and Philipp Koehn	6724
<i>Interactive Machine Comprehension with Dynamic Knowledge Graphs</i>	
Xingdi Yuan	6734

<i>Residual Adapters for Parameter-Efficient ASR Adaptation to Atypical and Accented Speech</i> Katrin Tomanek, Vicky Zayats, Dirk Padfield, Kara Vaillancourt and Fadi Biadisy	6751
<i>Visual News: Benchmark and Challenges in News Image Captioning</i> Fuxiao Liu, Yinghan Wang, Tianlu Wang and Vicente Ordonez.....	6761
<i>Integrating Visuospatial, Linguistic, and Commonsense Structure into Story Visualization</i> Adyasha Maharana and Mohit Bansal	6772
<i>VideoCLIP: Contrastive Pre-training for Zero-shot Video-Text Understanding</i> Hu Xu, Gargi Ghosh, Po-Yao Huang, Dmytro Okhonko, Armen Aghajanyan, Florian Metze, Luke Zettlemoyer and Christoph Feichtenhofer.....	6787
<i>NewsCLIPpings: Automatic Generation of Out-of-Context Multimodal Media</i> Grace Luo, Trevor Darrell and Anna Rohrbach	6801
<i>Powering Comparative Classification with Sentiment Analysis via Domain Adaptive Knowledge Transfer</i> Zeyu Li, Yilong Qin, Zihan Liu and Wei Wang	6818
<i>Tribrid: Stance Classification with Neural Inconsistency Detection</i> Song Yang and Jacopo Urbani	6831
<i>SYSML: StYlometry with Structure and Multitask Learning: Implications for Darknet Forum Migrant Analysis</i> Pranav Maneriker, Yuntian He and Srinivasan Parthasarathy	6844
<i>Few-Shot Emotion Recognition in Conversation with Sequential Prototypical Networks</i> Gaël Guibon, Matthieu Labeau, H�el�ene Flamein, Luce Lefeuvre and Chlo�e Clavel	6858
<i>CLASSIC: Continual and Contrastive Learning of Aspect Sentiment Classification Tasks</i> Zixuan Ke, Bing Liu, Hu Xu and Lei Shu	6871
<i>Implicit Sentiment Analysis with Event-centered Text Representation</i> Deyu Zhou, Jianan Wang, Linhai Zhang and Yulan He	6884
<i>SimCSE: Simple Contrastive Learning of Sentence Embeddings</i> Tianyu Gao, Xingcheng Yao and Danqi Chen	6894
<i>When is Wall a Pared and when a Muro?: Extracting Rules Governing Lexical Selection</i> Aditi Chaudhary, Kayo Yin, Antonios Anastasopoulos and Graham Neubig	6911
<i>Aligning Actions Across Recipe Graphs</i> Lucia Donatelli, Theresa Schmidt, Debanjali Biswas, Arne K�ohn, Fangzhou Zhai and Alexander Koller.....	6930
<i>Generating Datasets with Pretrained Language Models</i> Timo Schick and Hinrich Sch�utze	6943
<i>Continuous Entailment Patterns for Lexical Inference in Context</i> Martin Schmitt and Hinrich Sch�utze.....	6952
<i>Numeracy enhances the Literacy of Language Models</i> Avijit Thawani, Jay Pujara and Filip Ilievski.....	6960

<i>Students Who Study Together Learn Better: On the Importance of Collective Knowledge Distillation for Domain Transfer in Fact Verification</i>	
Mitch Paul Mithun, Sandeep Sunawal and Mihai Surdeanu	6968
<i>MultiEURLEX - A multi-lingual and multi-label legal document classification dataset for zero-shot cross-lingual transfer</i>	
Ilias Chalkidis, Manos Fergadiotis and Ion Androutsopoulos	6974
<i>Joint Passage Ranking for Diverse Multi-Answer Retrieval</i>	
Sewon Min, Kenton Lee, Ming-Wei Chang, Kristina Toutanova and Hannaneh Hajishirzi	6997
<i>Generative Context Pair Selection for Multi-hop Question Answering</i>	
Dheeru Dua, Cicero Nogueira dos Santos, Patrick Ng, Ben Athiwaratkun, Bing Xiang, Matt Gardner and Sameer Singh	7009
<i>Synthetic Data Augmentation for Zero-Shot Cross-Lingual Question Answering</i>	
Arij riabi, Thomas Scialom, Rachel Keraron, Benoît Sagot, Djamé Seddah and Jacopo Staiano	7016
<i>Have You Seen That Number? Investigating Extrapolation in Question Answering Models</i>	
Jeonghwan Kim, Giwon Hong, Kyung-min Kim, Junmo Kang and Sung-Hyon Myaeng	7031
<i>Surface Form Competition: Why the Highest Probability Answer Isn't Always Right</i>	
Ari Holtzman, Peter West, Vered Shwartz, Yejin Choi and Luke Zettlemoyer	7038
<i>Entity-Based Knowledge Conflicts in Question Answering</i>	
Shayne Longpre, Kartik Perisetla, Anthony Chen, Nikhil Ramesh, Chris DuBois and Sameer Singh	7052
<i>Back-Training excels Self-Training at Unsupervised Domain Adaptation of Question Generation and Passage Retrieval</i>	
Devang Kulshreshtha, Robert Belfer, Iulian Vlad Serban and Siva Reddy	7064
<i>DWUG: A large Resource of Diachronic Word Usage Graphs in Four Languages</i>	
Dominik Schlechtweg, Nina Tahmasebi, Simon Hengchen, Haim Dubossarsky and Barbara McGillivray	7079
<i>I Wish I Would Have Loved This One, But I Didn't – A Multilingual Dataset for Counterfactual Detection in Product Review</i>	
James O'Neill, Polina Rozenshtein, Ryuichi Kiryo, Motoko Kubota and Danushka Bollegala	7092
<i>Does It Capture STEL? A Modular, Similarity-based Linguistic Style Evaluation Framework</i>	
Anna Wegmann and Dong Nguyen	7109
<i>Evaluating the Morphosyntactic Well-formedness of Generated Texts</i>	
Adithya Pratapa, Antonios Anastasopoulos, Shruti Rijhwani, Aditi Chaudhary, David R. Mortensen, Graham Neubig and Yulia Tsvetkov	7131
<i>AM2iCo: Evaluating Word Meaning in Context across Low-Resource Languages with Adversarial Examples</i>	
Qianchu Liu, Edoardo Maria Ponti, Diana McCarthy, Ivan Vulić and Anna Korhonen	7151
<i>CrossFit: A Few-shot Learning Challenge for Cross-task Generalization in NLP</i>	
Qinyuan Ye, Bill Yuchen Lin and Xiang Ren	7163

<i>On the Influence of Masking Policies in Intermediate Pre-training</i>	
Qinyuan Ye, Belinda Z. Li, Sinong Wang, Benjamin Bolte, Hao Ma, Wen-tau Yih, Xiang Ren and Madian Khabsa	7190
<i>ValNorm Quantifies Semantics to Reveal Consistent Valence Biases Across Languages and Over Centuries</i>	
Autumn Toney and Aylin Caliskan	7203
<i>Perturbation CheckLists for Evaluating NLG Evaluation Metrics</i>	
Ananya B. Sai, Tanay Dixit, Dev Yashpal Sheth, Sreyas Mohan and Mitesh M. Khapra	7219
<i>Robust Open-Vocabulary Translation from Visual Text Representations</i>	
Elizabeth Salesky, David Etter and Matt Post	7235
<i>Don't Go Far Off: An Empirical Study on Neural Poetry Translation</i>	
Tuhin Chakrabarty, Arkadiy Saakyan and Smaranda Muresan	7253
<i>Improving Multilingual Translation by Representation and Gradient Regularization</i>	
Yilin Yang, Akiko Eriguchi, Alexandre Muzio, Prasad Tadepalli, Stefan Lee and Hany Hassan	7266
<i>Learning Kernel-Smoothed Machine Translation with Retrieved Examples</i>	
Qingnan Jiang, Mingxuan Wang, Jun Cao, Shanbo Cheng, Shujian Huang and Lei Li	7280
<i>Uncertainty-Aware Balancing for Multilingual and Multi-Domain Neural Machine Translation Training</i>	
Minghao Wu, Yitong Li, Meng Zhang, Liangyou Li, Gholamreza Haffari and Qun Liu	7291
<i>Universal Simultaneous Machine Translation with Mixture-of-Experts Wait-k Policy</i>	
Shaolei Zhang and Yang Feng	7306
<i>How much coffee was consumed during EMNLP 2019? Fermi Problems: A New Reasoning Challenge for AI</i>	
Ashwin Kalyan, Abhinav Kumar, Arjun Chandrasekaran, Ashish Sabharwal and Peter Clark .	7318
<i>Will this Question be Answered? Question Filtering via Answer Model Distillation for Efficient Question Answering</i>	
Siddhant Garg and Alessandro Moschitti	7329
<i>Learning with Instance Bundles for Reading Comprehension</i>	
Dheeru Dua, Pradeep Dasigi, Sameer Singh and Matt Gardner	7347
<i>Explaining Answers with Entailment Trees</i>	
Bhavana Dalvi, Peter Jansen, Oyvind Tafjord, Zhengnan Xie, Hannah Smith, Leighanna Pipatanangkura and Peter Clark	7358
<i>SituatedQA: Incorporating Extra-Linguistic Contexts into QA</i>	
Michael Zhang and Eunsol Choi	7371
<i>ConvAbuse: Data, Analysis, and Benchmarks for Nuanced Detection in Conversational AI</i>	
Amanda Cercas Curry, Gavin Abercrombie and Verena Rieser	7388
<i>Conversational Multi-Hop Reasoning with Neural Commonsense Knowledge and Symbolic Logic Rules</i>	
Forough Arabshahi, Jennifer Lee, Antoine Bosselut, Yejin Choi and Tom Mitchell	7404
<i>Towards Automatic Evaluation of Dialog Systems: A Model-Free Off-Policy Evaluation Approach</i>	
Haoming Jiang, Bo Dai, Mengjiao Yang, Tuo Zhao and Wei Wei	7419

<i>Continual Learning in Task-Oriented Dialogue Systems</i>	
Andrea Madotto, Zhaojiang Lin, Zhenpeng Zhou, Seungwhan Moon, Paul Crook, Bing Liu, Zhou Yu, Eunjoon Cho, Pascale Fung and Zhiguang Wang	7452
<i>Multilingual and Cross-Lingual Intent Detection from Spoken Data</i>	
Daniela Gerz, Pei-Hao Su, Razvan Kusztos, Avishek Mondal, Michał Lis, Eshan Singhal, Nikola Mrkšić, Tsung-Hsien Wen and Ivan Vulić	7468
<i>Investigating Robustness of Dialog Models to Popular Figurative Language Constructs</i>	
Harsh Jhamtani, Varun Gangal, Eduard Hovy and Taylor Berg-Kirkpatrick	7476
<i>Effective Sequence-to-Sequence Dialogue State Tracking</i>	
Jeffrey Zhao, Mahdis Mahdih, Ye Zhang, Yuan Cao and Yonghui Wu	7486
<i>MS²: Multi-Document Summarization of Medical Studies</i>	
Jay DeYoung, Iz Beltagy, Madeleine van Zuylen, Bailey Kuehl and Lucy Wang	7494
<i>CLIPScore: A Reference-free Evaluation Metric for Image Captioning</i>	
Jack Hessel, Ari Holtzman, Maxwell Forbes, Ronan Le Bras and Yejin Choi	7514
<i>On the Challenges of Evaluating Compositional Explanations in Multi-Hop Inference: Relevance, Completeness, and Expert Ratings</i>	
Peter Jansen, Kelly J. Smith, Dan Moreno and Huitzililn Ortiz	7529
<i>ESTER: A Machine Reading Comprehension Dataset for Reasoning about Event Semantic Relations</i>	
Rujun Han, I-Hung Hsu, Jiao Sun, Julia Baylon, Qiang Ning, Dan Roth and Nanyun Peng	7543
<i>RICA: Evaluating Robust Inference Capabilities Based on Commonsense Axioms</i>	
Pei Zhou, Rahul Khanna, Seyeon Lee, Bill Yuchen Lin, Daniel Ho, Jay Pujara and Xiang Ren	7560
<i>Compression, Transduction, and Creation: A Unified Framework for Evaluating Natural Language Generation</i>	
Mingkai Deng, Bowen Tan, Zhengzhong Liu, Eric Xing and Zhiting Hu	7580
<i>MATE: Multi-view Attention for Table Transformer Efficiency</i>	
Julian Eisenschlos, Maharshi Gor, Thomas Müller and William Cohen	7606
<i>Learning with Different Amounts of Annotation: From Zero to Many Labels</i>	
Shujian Zhang, Chengyue Gong and Eunsol Choi	7620
<i>When Attention Meets Fast Recurrence: Training Language Models with Reduced Compute</i>	
Tao Lei	7633
<i>Universal-KD: Attention-based Output-Grounded Intermediate Layer Knowledge Distillation</i>	
Yimeng Wu, Mehdi Rezagholizadeh, Abbas Ghaddar, Md Akmal Haidar and Ali Ghodsi	7649
<i>Highly Parallel Autoregressive Entity Linking with Discriminative Correction</i>	
Nicola De Cao, Wilker Aziz and Ivan Titov	7662
<i>Word-Level Coreference Resolution</i>	
Vladimir Dobrovolskii	7670
<i>A Secure and Efficient Federated Learning Framework for NLP</i>	
Chenghong Wang, Jieren Deng, Xianrui Meng, Yijue Wang, Ji Li, Sheng Lin, Shuo Han, Fei Miao, Sanguthevar Rajasekaran and Caiwen Ding	7676

<i>Controllable Semantic Parsing via Retrieval Augmentation</i>	
Panupong Pasupat, Yuan Zhang and Kelvin Guu	7683
<i>Constrained Language Models Yield Few-Shot Semantic Parsers</i>	
Richard Shin, Christopher Lin, Sam Thomson, Charles Chen, Subhro Roy, Emmanouil Antonios Platanios, Adam Pauls, Dan Klein, Jason Eisner and Benjamin Van Durme	7699
<i>ExplaGraphs: An Explanation Graph Generation Task for Structured Commonsense Reasoning</i>	
Swarnadeep Saha, Prateek Yadav, Lisa Bauer and Mohit Bansal	7716
<i>Connect-the-Dots: Bridging Semantics between Words and Definitions via Aligning Word Sense Inventories</i>	
Wenlin Yao, Xiaoman Pan, Lifeng Jin, Jianshu Chen, Dian Yu and Dong Yu	7741
<i>LM-Critic: Language Models for Unsupervised Grammatical Error Correction</i>	
Michihiro Yasunaga, Jure Leskovec and Percy Liang	7752
<i>Language-agnostic Representation from Multilingual Sentence Encoders for Cross-lingual Similarity Estimation</i>	
Nattapong Tiyajamorn, Tomoyuki Kajiwara, Yuki Arase and Makoto Onizuka	7764
<i>Classifying Dyads for Militarized Conflict Analysis</i>	
Niklas Stoehr, Lucas Torroba Hennigen, Samin Ahabab, Robert West and Ryan Cotterell	7775
<i>Point-of-Interest Type Prediction using Text and Images</i>	
Danae Sánchez Villegas and Nikolaos Aletras	7785
<i>Come hither or go away? Recognising pre-electoral coalition signals in the news</i>	
Ines Rehbein, Simone Paolo Ponzetto, Anna Adendorf, Oke Bahnsen, Lukas Stoetzer and Heiner Stuckenschmidt	7798
<i>#HowYouTagTweets: Learning User Hashtagging Preferences via Personalized Topic Attention</i>	
Yuji Zhang, Yubo Zhang, Chunpu Xu, Jing Li, Ziyang Jiang and Baolin Peng	7811
<i>Learning Neural Templates for Recommender Dialogue System</i>	
Zujie Liang, Huang Hu, Can Xu, Jian Miao, Yingying He, Yining Chen, Xiubo Geng, Fan Liang and Daxin Jiang	7821
<i>Proxy Indicators for the Quality of Open-domain Dialogues</i>	
Rostislav Nedelchev, Jens Lehmann and Ricardo Usbeck	7834
<i>Q²: Evaluating Factual Consistency in Knowledge-Grounded Dialogues via Question Generation and Question Answering</i>	
Or Honovich, Leshem Choshen, Roei Aharoni, Ella Neeman, Idan Szpektor and Omri Abend	7856
<i>Knowledge-Aware Graph-Enhanced GPT-2 for Dialogue State Tracking</i>	
Weizhe Lin, Bo-Hsiang Tseng and Bill Byrne	7871
<i>A Collaborative Multi-agent Reinforcement Learning Framework for Dialog Action Decomposition</i>	
Huimin Wang and Kam-Fai Wong	7882
<i>Zero-Shot Dialogue State Tracking via Cross-Task Transfer</i>	
Zhaojiang Lin, Bing Liu, Andrea Madotto, Seungwhan Moon, Zhenpeng Zhou, Paul Crook, Zhiguang Wang, Zhou Yu, Eunjoon Cho, Rajen Subba and Pascale Fung	7890

<i>Uncertainty Measures in Neural Belief Tracking and the Effects on Dialogue Policy Performance</i> Carel van Niekerk, Andrey Malinin, Christian Geishausser, Michael Heck, Hsien-chin Lin, Nurul Lubis, Shutong Feng and Milica Gasic	7901
<i>Dynamic Forecasting of Conversation Derailment</i> Yova Kementchedjhieva and Anders Søgaard	7915
<i>A Semantic Filter Based on Relations for Knowledge Graph Completion</i> Zongwei Liang, Junan Yang, Hui Liu and Keju Huang	7920
<i>AdapterDrop: On the Efficiency of Adapters in Transformers</i> Andreas Rücklé, Gregor Geigle, Max Glockner, Tilman Beck, Jonas Pfeiffer, Nils Reimers and Iryna Gurevych	7930
<i>Understanding and Overcoming the Challenges of Efficient Transformer Quantization</i> Yelysei Bondarenko, Markus Nagel and tijmen Blankevoort	7947
<i>CAPE: Context-Aware Private Embeddings for Private Language Learning</i> Richard Plant, Dimitra Gkatzia and Valerio Giuffrida	7970
<i>Text Detoxification using Large Pre-trained Neural Models</i> David Dale, Anton Voronov, Daryna Dementieva, Varvara Logacheva, Olga Kozlova, Nikita Semenov and Alexander Panchenko	7979
<i>Document-Level Text Simplification: Dataset, Criteria and Baseline</i> Renliang Sun, Hanqi Jin and Xiaojun Wan	7997
<i>A Bag of Tricks for Dialogue Summarization</i> Muhammad Khalifa, Miguel Ballesteros and Kathleen McKeown	8014
<i>Paraphrasing Compound Nominalizations</i> John Lee, Ho Hung Lim and Carol Webster	8023
<i>Data-QuestEval: A Referenceless Metric for Data-to-Text Semantic Evaluation</i> Clement Rebuffel, Thomas Scialom, Laure Soulier, Benjamin Piwowarski, Sylvain Lamprier, Jacopo Staiano, Geoffrey Scoutheeten and patrick Gallinari	8029
<i>Low-Rank Subspaces for Unsupervised Entity Linking</i> Akhil Arora, Alberto Garcia-Duran and Robert West	8037
<i>TDEER: An Efficient Translating Decoding Schema for Joint Extraction of Entities and Relations</i> Xianming Li, Xiaotian Luo, Chenghao Dong, Daichuan Yang, Beidi Luan and Zhen He	8055
<i>Extracting Event Temporal Relations via Hyperbolic Geometry</i> Xingwei Tan, Gabriele Pergola and Yulan He	8065
<i>Honey or Poison? Solving the Trigger Curse in Few-shot Event Detection via Causal Intervention</i> Jiawei Chen, Hongyu Lin, Xianpei Han and Le Sun	8078
<i>Back to the Basics: A Quantitative Analysis of Statistical and Graph-Based Term Weighting Schemes for Keyword Extraction</i> Asahi Ushio, Federico Liberatore and Jose Camacho-Collados	8089
<i>Time-dependent Entity Embedding is not All You Need: A Re-evaluation of Temporal Knowledge Graph Completion Models under a Unified Framework</i> Zhen Han, Gengyuan Zhang, Yunpu Ma and Volker Tresp	8104

<i>Matching-oriented Embedding Quantization For Ad-hoc Retrieval</i>	
Shitao Xiao, Zheng Liu, Yingxia Shao, Defu Lian and Xing Xie	8119
<i>Efficient Mind-Map Generation via Sequence-to-Graph and Reinforced Graph Refinement</i>	
Mengting Hu, Honglei Guo, Shiwang Zhao, Hang Gao and Zhong Su	8130
<i>Deep Attention Diffusion Graph Neural Networks for Text Classification</i>	
Yonghao Liu, Renchu Guan, fausto giunchiglia, Yanchun Liang and Xiaoyue Feng	8142
<i>Balancing Methods for Multi-label Text Classification with Long-Tailed Class Distribution</i>	
Yi Huang, Buse Giledereli, Abdullatif Köksal, Arzucan Özgür and Elif Ozkirimli	8153
<i>Bayesian Topic Regression for Causal Inference</i>	
Maximilian Ahrens, Julian Ashwin, Jan-Peter Calliess and Vu Nguyen	8162
<i>Enjoy the Saliency: Towards Better Transformer-based Faithful Explanations with Word Saliency</i>	
George Chrysostomou and Nikolaos Aletras	8189
<i>What’s in Your Head? Emergent Behaviour in Multi-Task Transformer Models</i>	
Mor Geva, Uri Katz, Aviv Ben-Arie and Jonathan Berant	8201
<i>Don’t Search for a Search Method — Simple Heuristics Suffice for Adversarial Text Attacks</i>	
Nathaniel Berger, Stefan Riezler, Sebastian Ebert and Artem Sokolov	8216
<i>Adversarial Attacks on Knowledge Graph Embeddings via Instance Attribution Methods</i>	
Peru Bhardwaj, John Kelleher, Luca Costabello and Declan O’Sullivan	8225
<i>Locke’s Holiday: Belief Bias in Machine Reading</i>	
Anders Søgaard	8240
<i>Sequence Length is a Domain: Length-based Overfitting in Transformer Models</i>	
Dusan Varis and Ondřej Bojar	8246
<i>Contrasting Human- and Machine-Generated Word-Level Adversarial Examples for Text Classification</i>	
Maximilian Mozes, Max Bartolo, Pontus Stenetorp, Bennett Kleinberg and Lewis Griffin	8258
<i>Is Information Density Uniform in Task-Oriented Dialogues?</i>	
Mario Giulianelli, Arabella Sinclair and Raquel Fernández	8271
<i>On Homophony and Rényi Entropy</i>	
Tiago Pimentel, Clara Meister, Simone Teufel and Ryan Cotterell	8284
<i>Synthetic Textual Features for the Large-Scale Detection of Basic-level Categories in English and Mandarin</i>	
Yiwen Chen and Simone Teufel	8294
<i>TimeTraveler: Reinforcement Learning for Temporal Knowledge Graph Forecasting</i>	
Haohai Sun, Jialun Zhong, Yunpu Ma, Zhen Han and Kun He	8306
<i>Code-switched inspired losses for spoken dialog representations</i>	
Pierre Colombo, Emile Chapuis, Matthieu Labeau and Chloé Clavel	8320
<i>BiQUE: Biquaternionic Embeddings of Knowledge Graphs</i>	
Jia Guo and Stanley Kok	8338

<i>Learning Neural Ordinary Equations for Forecasting Future Links on Temporal Knowledge Graphs</i> Zhen Han, Zifeng Ding, Yunpu Ma, Yujia Gu and Volker Tresp	8352
<i>RAP: Robustness-Aware Perturbations for Defending against Backdoor Attacks on NLP Models</i> Wenkai Yang, Yankai Lin, Peng Li, Jie Zhou and Xu Sun	8365
<i>FAME: Feature-Based Adversarial Meta-Embeddings for Robust Input Representations</i> Lukas Lange, Heike Adel, Jannik Strötgen and Dietrich Klakow	8382
<i>A Strong Baseline for Query Efficient Attacks in a Black Box Setting</i> Rishabh Maheshwary, Saket Maheshwary and Vikram Pudi	8396
<i>Machine Translation Decoding beyond Beam Search</i> Rémi Leblond, Jean-Baptiste Alayrac, Laurent Sifre, Miruna Pîslar, Lespiau Jean-Baptiste, Ioannis Antonoglou, Karen Simonyan and Oriol Vinyals	8410
<i>Document Graph for Neural Machine Translation</i> Mingzhou Xu, Liangyou Li, Derek F. Wong, Qun Liu and Lidia S. Chao	8435
<i>An Empirical Investigation of Word Alignment Supervision for Zero-Shot Multilingual Neural Machine Translation</i> Alessandro Raganato, Raúl Vázquez, Mathias Creutz and Jörg Tiedemann	8449
<i>Graph Algorithms for Multiparallel Word Alignment</i> Ayyoob ImaniGooghari, Masoud Jalili Sabet, Lutfi Kerem Senel, Philipp Dufter, François Yvon and Hinrich Schütze	8457
<i>Improving the Quality Trade-Off for Neural Machine Translation Multi-Domain Adaptation</i> Eva Hasler, Tobias Domhan, Jonay Trenous, Ke Tran, Bill Byrne and Felix Hieber	8470
<i>Language Modeling, Lexical Translation, Reordering: The Training Process of NMT through the Lens of Classical SMT</i> Elena Voita, Rico Sennrich and Ivan Titov	8478
<i>Effective Fine-Tuning Methods for Cross-lingual Adaptation</i> Tao Yu and Shafiq Joty	8492
<i>Rethinking Data Augmentation for Low-Resource Neural Machine Translation: A Multi-Task Learning Approach</i> Víctor M. Sánchez-Cartagena, Miquel Esplà-Gomis, Juan Antonio Pérez-Ortiz and Felipe Sánchez-Martínez	8502
<i>Wino-X: Multilingual Winograd Schemas for Commonsense Reasoning and Coreference Resolution</i> Denis Emelin and Rico Sennrich	8517
<i>One Source, Two Targets: Challenges and Rewards of Dual Decoding</i> Jitao Xu and François Yvon	8533
<i>Discrete and Soft Prompting for Multilingual Models</i> Mengjie Zhao and Hinrich Schütze	8547
<i>Vision Matters When It Should: Sanity Checking Multimodal Machine Translation Models</i> Jiaoda Li, Duygu Ataman and Rico Sennrich	8556
<i>Efficient Inference for Multilingual Neural Machine Translation</i> Alexandre Berard, Dain Lee, Stéphane Clinchant, Kweonwoo Jung and Vassilina Nikoulina ..	8563

<i>Role of Language Relatedness in Multilingual Fine-tuning of Language Models: A Case Study in Indo-Aryan Languages</i>	
Tejas Dhamecha, Rudra Murthy, Samarth Bharadwaj, Karthik Sankaranarayanan and Pushpak Bhat-tacharyya	8584
<i>Comparing Feature-Engineering and Feature-Learning Approaches for Multilingual Translationese Clas-sification</i>	
Daria Pylypenko, Kwabena Amponsah-Kaakyire, Koel Dutta Chowdhury, Josef van Genabith and Cristina España-Bonet	8596
<i>Multi-Sentence Resampling: A Simple Approach to Alleviate Dataset Length Bias and Beam-Search Degradation</i>	
Ivan Provilkov and Andrey Malinin	8612
<i>Cross-Policy Compliance Detection via Question Answering</i>	
Marzieh Saeidi, Majid Yazdani and Andreas Vlachos	8622
<i>Meta-LMTC: Meta-Learning for Large-Scale Multi-Label Text Classification</i>	
Ran Wang, Xi’ao Su, Siyu Long, Xinyu Dai, Shujian Huang and Jiajun Chen	8633
<i>Unsupervised Multi-View Post-OCR Error Correction With Language Models</i>	
Harsh Gupta, Luciano Del Corro, Samuel Broscheit, Johannes Hoffart and Eliot Brenner	8647
<i>Parallel Refinements for Lexically Constrained Text Generation with BART</i>	
Xingwei He	8653
<i>BERT-Beta: A Proactive Probabilistic Approach to Text Moderation</i>	
Fei Tan, Yifan Hu, Kevin Yen and Changwei Hu	8667
<i>STaCK: Sentence Ordering with Temporal Commonsense Knowledge</i>	
Deepanway Ghosal, Navonil Majumder, Rada Mihalcea and Soujanya Poria	8676
<i>Preventing Author Profiling through Zero-Shot Multilingual Back-Translation</i>	
David Adelani, Miaoran Zhang, Xiaoyu Shen, Ali Davody, Thomas Kleinbauer and Dietrich Klakow	8687
<i>CodeT5: Identifier-aware Unified Pre-trained Encoder-Decoder Models for Code Understanding and Generation</i>	
Yue Wang, Weishi Wang, Shafiq Joty and Steven C.H. Hoi	8696
<i>Detect and Classify – Joint Span Detection and Classification for Health Outcomes</i>	
Micheal Abaho, Danushka Bollegala, Paula Williamson and Susanna Dodd	8709
<i>Multi-Class Grammatical Error Detection for Correction: A Tale of Two Systems</i>	
Zheng Yuan, Shiva Taslimipoor, Christopher Davis and Christopher Bryant	8722
<i>Towards Zero-shot Commonsense Reasoning with Self-supervised Refinement of Language Models</i>	
Tassilo Klein and Moin Nabi	8737
<i>To Share or not to Share: Predicting Sets of Sources for Model Transfer Learning</i>	
Lukas Lange, Jannik Strötgen, Heike Adel and Dietrich Klakow	8744

<i>Self-Supervised Detection of Contextual Synonyms in a Multi-Class Setting: Phenotype Annotation Use Case</i>	
Jingqing Zhang, Luis Bolanos Trujillo, Tong Li, Ashwani Tanwar, Guilherme Freire, Xian Yang, Julia Ive, Vibhor Gupta and Yike Guo	8754
<i>ClauseRec: A Clause Recommendation Framework for AI-aided Contract Authoring</i>	
Vinay Aggarwal, Aparna Garimella, Balaji Vasani Srinivasan, Anandhavelu N and Rajiv Jain .	8770
<i>Finnish Dialect Identification: The Effect of Audio and Text</i>	
Mika Hämmäläinen, Khalid Alnajjar, Niko Partanen and Jack Rueter	8777
<i>English Machine Reading Comprehension Datasets: A Survey</i>	
Daria Dzendzik, Jennifer Foster and Carl Vogel	8784
<i>Expanding End-to-End Question Answering on Differentiable Knowledge Graphs with Intersection</i>	
Priyanka Sen, Armin Oliya and Amir Saffari	8805
<i>Structured Context and High-Coverage Grammar for Conversational Question Answering over Knowledge Graphs</i>	
Pierre Marion, Pawel Nowak and Francesco Piccinno	8813
<i>Improving Question Answering Model Robustness with Synthetic Adversarial Data Generation</i>	
Max Bartolo, Tristan Thrush, Robin Jia, Sebastian Riedel, Pontus Stenetorp and Douwe Kiela	8830
<i>BeliefBank: Adding Memory to a Pre-Trained Language Model for a Systematic Notion of Belief</i>	
Nora Kassner, Oyvind Tafjord, Hinrich Schütze and Peter Clark	8849
<i>MLEC-QA: A Chinese Multi-Choice Biomedical Question Answering Dataset</i>	
Jing Li, Shangping Zhong and Kaizhi Chen	8862
<i>IndoNLG: Benchmark and Resources for Evaluating Indonesian Natural Language Generation</i>	
Samuel Cahyawijaya, Genta Indra Winata, Bryan Wilie, Karissa Vincentio, Xiaohong Li, Adhiguna Kuncoro, Sebastian Ruder, Zhi Yuan Lim, Syafri Bahar, Masayu Khodra, Ayu Purwarianti and Pascale Fung	8875
<i>Is Multi-Hop Reasoning Really Explainable? Towards Benchmarking Reasoning Interpretability</i>	
Xin Lv, Yixin Cao, Lei Hou, Juanzi Li, Zhiyuan Liu, Yichi Zhang and Zelin Dai	8899
<i>Global Explainability of BERT-Based Evaluation Metrics by Disentangling along Linguistic Factors</i>	
Marvin Kaster, Wei Zhao and Steffen Eger	8912
<i>Exploring Underexplored Limitations of Cross-Domain Text-to-SQL Generalization</i>	
Yujian Gan, Xinyun Chen and Matthew Purver	8926
<i>What happens if you treat ordinal ratings as interval data? Human evaluations in NLP are even more under-powered than you think</i>	
David M. Howcroft and Verena Rieser	8932
<i>NeuTral Rewriter: A Rule-Based and Neural Approach to Automatic Rewriting into Gender Neutral Alternatives</i>	
Eva Vanmassenhove, Chris Emmery and Dimitar Shterionov	8940
<i>Benchmarking Commonsense Knowledge Base Population with an Effective Evaluation Dataset</i>	
Tianqing Fang, Weiqi Wang, Sehyun Choi, Shibo Hao, Hongming Zhang, Yangqiu Song and Bin He	8949

<i>Enhancing the Context Representation in Similarity-based Word Sense Disambiguation</i>	
Ming Wang, Jianzhang Zhang and Yinglin Wang	8965
<i>Data Augmentation with Hierarchical SQL-to-Question Generation for Cross-domain Text-to-SQL Parsing</i>	
Kun Wu, Lijie Wang, Zhenghua Li, Ao Zhang, Xinyan Xiao, Hua Wu, Min Zhang and Haifeng Wang	8974
<i>SPARQLing Database Queries from Intermediate Question Decompositions</i>	
Irina Saparina and Anton Osokin	8984
<i>Time-aware Graph Neural Network for Entity Alignment between Temporal Knowledge Graphs</i>	
Chengjin Xu, Fenglong Su and Jens Lehmann	8999
<i>Cross-Domain Label-Adaptive Stance Detection</i>	
Momchil Hardalov, Arnav Arora, Preslav Nakov and Isabelle Augenstein	9011
<i>Text AutoAugment: Learning Compositional Augmentation Policy for Text Classification</i>	
Shuhuai Ren, Jinchao Zhang, Lei Li, Xu Sun and Jie Zhou	9029
<i>Distilling Relation Embeddings from Pretrained Language Models</i>	
Asahi Ushio, Jose Camacho-Collados and Steven Schockaert	9044
<i>Avoiding Inference Heuristics in Few-shot Prompt-based Finetuning</i>	
Prasetya Utama, Nafise Sadat Moosavi, Victor Sanh and Iryna Gurevych	9063
<i>A Differentiable Relaxation of Graph Segmentation and Alignment for AMR Parsing</i>	
Chunchuan Lyu, Shay B. Cohen and Ivan Titov	9075
<i>Integrating Personalized PageRank into Neural Word Sense Disambiguation</i>	
Ahmed El Sheikh, Michele Bevilacqua and Roberto Navigli	9092
<i>Cross-lingual Sentence Embedding using Multi-Task Learning</i>	
Koustava Goswami, Sourav Dutta, Haytham Assem, Theodorus Fransen and John P. McCrae	9099
<i>NB-MLM: Efficient Domain Adaptation of Masked Language Models for Sentiment Analysis</i>	
Nikolay Arefyev, Dmitrii Kharchev and Artem Shelmanov	9114
<i>Revisiting Self-training for Few-shot Learning of Language Model</i>	
Yiming Chen, Yan Zhang, Chen Zhang, Grandee Lee, Ran Cheng and Haizhou Li	9125
<i>Bridging Perception, Memory, and Inference through Semantic Relations</i>	
Johanna Björklund, Adam Dahlgren Lindström and Frank Drewes	9136
<i>Unimodal and Crossmodal Refinement Network for Multimodal Sequence Fusion</i>	
Xiaobao Guo, Adams Kong, Huan Zhou, Xianfeng Wang and Min Wang	9143
<i>YASO: A Targeted Sentiment Analysis Evaluation Dataset for Open-Domain Reviews</i>	
Matan Orbach, Orith Toledo-Ronen, Artem Spector, Ranit Aharonov, Yoav Katz and Noam Slonim	9154
<i>An Empirical Study on Leveraging Position Embeddings for Target-oriented Opinion Words Extraction</i>	
Samuel Mensah, Kai Sun and Nikolaos Aletras	9174

<i>Improving Multimodal Fusion with Hierarchical Mutual Information Maximization for Multimodal Sentiment Analysis</i>	
Wei Han, Hui Chen and Soujanya Poria	9180
<i>BERT4GCN: Using BERT Intermediate Layers to Augment GCN for Aspect-based Sentiment Classification</i>	
Zeguan Xiao, Jiarun Wu, Qingliang Chen and Congjian Deng	9193
<i>Does Social Pressure Drive Persuasion in Online Fora?</i>	
Ayush Jain and Shashank Srivastava	9201
<i>Aspect Sentiment Quad Prediction as Paraphrase Generation</i>	
Wenxuan Zhang, Yang Deng, Xin Li, Yifei Yuan, Lidong Bing and Wai Lam	9209
<i>Cross-lingual Aspect-based Sentiment Analysis with Aspect Term Code-Switching</i>	
Wenxuan Zhang, Ruidan He, Haiyun Peng, Lidong Bing and Wai Lam	9220
<i>Towards Label-Agnostic Emotion Embeddings</i>	
Sven Buechel, Luise Modersohn and Udo Hahn	9231
<i>Collaborative Learning of Bidirectional Decoders for Unsupervised Text Style Transfer</i>	
Yun Ma, Yangbin Chen, Xudong Mao and Qing Li	9250
<i>Exploring Non-Autoregressive Text Style Transfer</i>	
Yun Ma and Qing Li	9267
<i>PASTE: A Tagging-Free Decoding Framework Using Pointer Networks for Aspect Sentiment Triplet Extraction</i>	
Rajdeep Mukherjee, Tapas Nayak, Yash Butala, Sourangshu Bhattacharya and Pawan Goyal	9279
<i>Adaptive Proposal Generation Network for Temporal Sentence Localization in Videos</i>	
Daizong Liu, Xiaoye Qu, Jianfeng Dong and Pan Zhou	9292
<i>Progressively Guide to Attend: An Iterative Alignment Framework for Temporal Sentence Grounding</i>	
Daizong Liu, Xiaoye Qu and Pan Zhou	9302
<i>Language Models are Few-Shot Butlers</i>	
Vincent Micheli and Francois Fleuret	9312
<i>R³Net: Relation-embedded Representation Reconstruction Network for Change Captioning</i>	
Yunbin Tu, Liang Li, Chenggang Yan, Shengxiang Gao and Zhengtao Yu	9319
<i>Looking for Confirmations: An Effective and Human-Like Visual Dialogue Strategy</i>	
Alberto Testoni and Raffaella Bernardi	9330
<i>A Unified Speaker Adaptation Approach for ASR</i>	
Yingzhu Zhao, Chongjia Ni, Cheung-Chi LEUNG, Shafiq Joty, Eng Siong Chng and Bin Ma	9339
<i>Caption Enriched Samples for Improving Hateful Memes Detection</i>	
Efrat Blaier, Itzik Malkiel and Lior Wolf	9350
<i>Sparsity and Sentence Structure in Encoder-Decoder Attention of Summarization Systems</i>	
Potsawee Manakul and Mark Gales	9359
<i>BARThez: a Skilled Pretrained French Sequence-to-Sequence Model</i>	
Moussa Kamal Eddine, Antoine Tixier and Michalis Vazirgiannis	9369

<i>ARMAN: Pre-training with Semantically Selecting and Reordering of Sentences for Persian Abstractive Summarization</i>	
Alireza Salemi, Emad Kebriaei, Ghazal Neisi Minaei and Azadeh Shakery	9391
<i>Models and Datasets for Cross-Lingual Summarisation</i>	
Laura Perez-Beltrachini and Mirella Lapata	9408
<i>Learning Opinion Summarizers by Selecting Informative Reviews</i>	
Arthur Bražinskas, Mirella Lapata and Ivan Titov	9424
<i>Enriching and Controlling Global Semantics for Text Summarization</i>	
Thong Nguyen, Anh Tuan Luu, Truc Lu and Tho Quan	9443
<i>Revisiting Tri-training of Dependency Parsers</i>	
Joachim Wagner and Jennifer Foster	9457
<i>Bridge to Target Domain by Prototypical Contrastive Learning and Label Confusion: Re-explore Zero-Shot Learning for Slot Filling</i>	
Liwen Wang, Xuefeng Li, Jiachi Liu, Keqing He, Yuanmeng Yan and Weiran Xu	9474
<i>Neuralizing Regular Expressions for Slot Filling</i>	
Chengyue Jiang, Zijian Jin and Kewei Tu	9481
<i>Causal Direction of Data Collection Matters: Implications of Causal and Anticausal Learning for NLP</i>	
Zhijing Jin, Julius von Kügelgen, Jingwei Ni, Tejas Vaidhya, Ayush Kaushal, Mrinmaya Sachan and Bernhard Schoelkopf	9499
<i>Raise a Child in Large Language Model: Towards Effective and Generalizable Fine-tuning</i>	
Runxin Xu, Fuli Luo, Zhiyuan Zhang, Chuanqi Tan, Baobao Chang, Songfang Huang and Fei Huang	9514
<i>Knowledge Graph Representation Learning using Ordinary Differential Equations</i>	
Mojtaba Nayyeri, Chengjin Xu, Franca Hoffmann, Mirza Mohtashim Alam, Jens Lehmann and sahar vahdati	9529
<i>KnowMAN: Weakly Supervised Multinomial Adversarial Networks</i>	
Luisa März, Ehsaneddin Asgari, Fabienne Braune, Franziska Zimmermann and Benjamin Roth	9549
<i>ONION: A Simple and Effective Defense Against Textual Backdoor Attacks</i>	
Fanchao Qi, Yangyi Chen, Mukai Li, Yuan Yao, Zhiyuan Liu and Maosong Sun	9558
<i>Value-aware Approximate Attention</i>	
Ankit Gupta and Jonathan Berant	9567
<i>Contrastive Domain Adaptation for Question Answering using Limited Text Corpora</i>	
Zhenrui Yue, Bernhard Kratzwald and Stefan Feuerriegel	9575
<i>Case-based Reasoning for Natural Language Queries over Knowledge Bases</i>	
Rajarshi Das, Manzil Zaheer, Dung Thai, Ameya Godbole, Ethan Perez, Jay Yoon Lee, Lizhen Tan, Lazaros Polymenakos and Andrew McCallum	9594
<i>Distantly-Supervised Dense Retrieval Enables Open-Domain Question Answering without Evidence Annotation</i>	
Chen Zhao, Chenyan Xiong, Jordan Boyd-Graber and Hal Daumé III	9612

<i>What's in a Name? Answer Equivalence For Open-Domain Question Answering</i>	
Chenglei Si, Chen Zhao and Jordan Boyd-Graber	9623
<i>Evaluation Paradigms in Question Answering</i>	
Pedro Rodriguez and Jordan Boyd-Graber	9630
<i>Numerical reasoning in machine reading comprehension tasks: are we there yet?</i>	
Hadeel Al-Negheimish, Pranava Madhyastha and Alessandra Russo	9643
<i>Set Generation Networks for End-to-End Knowledge Base Population</i>	
Dianbo Sui, Chenhao Wang, Yubo Chen, Kang Liu, Jun Zhao and Wei Bi	9650
<i>Knowing False Negatives: An Adversarial Training Method for Distantly Supervised Relation Extraction</i>	
Kailong Hao, Botao Yu and Wei Hu	9661
<i>Progressive Adversarial Learning for Bootstrapping: A Case Study on Entity Set Expansion</i>	
Lingyong Yan, Xianpei Han and Le Sun	9673
<i>Uncovering Main Causalities for Long-tailed Information Extraction</i>	
Guoshun Nan, Jiaqi Zeng, Rui Qiao, Zhijiang Guo and Wei Lu	9683
<i>Maximal Clique Based Non-Autoregressive Open Information Extraction</i>	
Bowen Yu, Yucheng Wang, Tingwen Liu, Hongsong Zhu, Limin Sun and Bin Wang	9696
<i>A Relation-Oriented Clustering Method for Open Relation Extraction</i>	
Jun Zhao, Tao Gui, Qi Zhang and Yaqian Zhou	9707
<i>Exploring Methods for Generating Feedback Comments for Writing Learning</i>	
Kazuaki Hanawa, Ryo Nagata and Kentaro Inui	9719
<i>A Role-Selected Sharing Network for Joint Machine-Human Chatting Handoff and Service Satisfaction Analysis</i>	
Jiawei Liu, Kaisong Song, Yangyang Kang, Guoxiu He, Zhuoren Jiang, Changlong Sun, Wei Lu and Xiaozhong Liu	9731
<i>Meta Distant Transfer Learning for Pre-trained Language Models</i>	
Chengyu Wang, Haojie Pan, Minghui Qiu, Jun Huang, Fei Yang and Yin Zhang	9742
<i>UniKER: A Unified Framework for Combining Embedding and Definite Horn Rule Reasoning for Knowledge Graph Inference</i>	
Kewei Cheng, Ziqing Yang, Ming Zhang and Yizhou Sun	9753
<i>Wasserstein Selective Transfer Learning for Cross-domain Text Mining</i>	
Lingyun Feng, Minghui Qiu, Yaliang Li, Haitao Zheng and Ying Shen	9772
<i>Jointly Learning to Repair Code and Generate Commit Message</i>	
Jiaqi Bai, Long Zhou, Ambrosio Blanco, Shujie Liu, Furu Wei, Ming Zhou and Zhoujun Li ..	9784
<i>Inflate and Shrink: Enriching and Reducing Interactions for Fast Text-Image Retrieval</i>	
Haoliang Liu, Tan Yu and Ping Li	9796
<i>On Pursuit of Designing Multi-modal Transformer for Video Grounding</i>	
Meng Cao, Long Chen, Mike Zheng Shou, Can Zhang and Yuexian Zou	9810
<i>COVR: A Test-Bed for Visually Grounded Compositional Generalization with Real Images</i>	
Ben Bogin, Shivanshu Gupta, Matt Gardner and Jonathan Berant	9824

<i>Vision-and-Language or Vision-for-Language? On Cross-Modal Influence in Multimodal Transformers</i> Stella Frank, Emanuele Bugliarello and Desmond Elliott	9847
<i>HypMix: Hyperbolic Interpolative Data Augmentation</i> Ramit Sawhney, Megh Thakkar, Shivam Agarwal, Di Jin, Diyi Yang and Lucie Flek	9858
<i>Integrating Deep Event-Level and Script-Level Information for Script Event Prediction</i> Long Bai, Saiping Guan, Jiafeng Guo, Zixuan Li, Xiaolong Jin and Xueqi Cheng	9869
<i>QA-Align: Representing Cross-Text Content Overlap by Aligning Question-Answer Propositions</i> Daniela Brook Weiss, Paul Roit, Ayal Klein, Ori Ernst and Ido Dagan	9879
<i>PICARD: Parsing Incrementally for Constrained Auto-Regressive Decoding from Language Models</i> Torsten Scholak, Nathan Schucher and Dzmitry Bahdanau	9895
<i>Exploiting Twitter as Source of Large Corpora of Weakly Similar Pairs for Semantic Sentence Embeddings</i> Marco Di Giovanni and Marco Brambilla	9902
<i>Guilt by Association: Emotion Intensities in Lexical Representations</i> Shahab Raji and Gerard de Melo	9911
<i>Using Sociolinguistic Variables to Reveal Changing Attitudes Towards Sexuality and Gender</i> Sky CH-Wang and David Jurgens	9918
<i>Identifying Morality Frames in Political Tweets using Relational Learning</i> Shamik Roy, Maria Leonor Pacheco and Dan Goldwasser	9939
<i>Measuring Sentence-Level and Aspect-Level (Un)certainly in Science Communications</i> Jiaxin Pei and David Jurgens	9959
<i>Assessing the Reliability of Word Embedding Gender Bias Measures</i> Yupei Du, Qixiang Fang and Dong Nguyen	10012
<i>Rumor Detection on Twitter with Claim-Guided Hierarchical Graph Attention Networks</i> Hongzhan Lin, Jing Ma, Mingfei Cheng, Zhiwei Yang, Liangliang Chen and Guang Chen ..	10035
<i>Learning Bill Similarity with Annotated and Augmented Corpora of Bills</i> Jiseon Kim, Elden Griggs, In Song Kim and Alice Oh.....	10048
<i>SWEAT: Scoring Polarization of Topics across Different Corpora</i> Federico Bianchi, Marco Marelli, Paolo Nicoli and Matteo Palmonari	10065
<i>"So You Think You're Funny?": Rating the Humour Quotient in Standup Comedy</i> Anirudh Mittal, Pranav Jeevan P, Prerak Gandhi, Diptesh Kanojia and Pushpak Bhattacharyya	10073
<i>"Was it "stated" or was it "claimed"?: How linguistic bias affects generative language models</i> Roma Patel and Ellie Pavlick	10080
<i>PAUSE: Positive and Annealed Unlabeled Sentence Embedding</i> Lele Cao, Emil Larsson, Vilhelm von Ehrenheim, Dhiana Deva Cavalcanti Rocha, Anna Martin and Sonja Horn	10096
<i>A Simple Geometric Method for Cross-Lingual Linguistic Transformations with Pre-trained Autoencoders</i> Maarten De Raedt, Frédéric Godin, Pieter Buteneers, Chris Develder and Thomas Demeester	10108

<i>An Information-Theoretic Characterization of Morphological Fusion</i>	
Neil Rathi, Michael Hahn and Richard Futrell	10115
<i>The Effect of Efficient Messaging and Input Variability on Neural-Agent Iterated Language Learning</i>	
Yuchen Lian, Arianna Bisazza and Tessa Verhoef	10121
<i>On Classifying whether Two Texts are on the Same Side of an Argument</i>	
Erik Körner, Gregor Wiedemann, Ahmad Dawar Hakimi, Gerhard Heyer and Martin Potthast	10130
<i>Chinese Opinion Role Labeling with Corpus Translation: A Pivot Study</i>	
Ranran Zhen, Rui Wang, Guohong Fu, Chengguo Lv and Meishan Zhang	10139
<i>MassiveSumm: a very large-scale, very multilingual, news summarisation dataset</i>	
Daniel Varab and Natalie Schluter	10150
<i>AUTOSUMM: Automatic Model Creation for Text Summarization</i>	
Sharmila Reddy Nangi, Atharv Tyagi, Jay Mundra, Sagnik Mukherjee, Raj Snehal, Niyati Chhaya and Aparna Garimella	10162
<i>Investigating the Helpfulness of Word-Level Quality Estimation for Post-Editing Machine Translation Output</i>	
Raksha Shenoy, Nico Herbig, Antonio Krüger and Josef van Genabith	10173
<i>UNKs Everywhere: Adapting Multilingual Language Models to New Scripts</i>	
Jonas Pfeiffer, Ivan Vulić, Iryna Gurevych and Sebastian Ruder	10186
<i>Neural Machine Translation Quality and Post-Editing Performance</i>	
Vilém Zouhar, Martin Popel, Ondřej Bojar and Aleš Tamchyna	10204
<i>XTREME-R: Towards More Challenging and Nuanced Multilingual Evaluation</i>	
Sebastian Ruder, Noah Constant, Jan Botha, Aditya Siddhant, Orhan Firat, Jinlan Fu, Pengfei Liu, Junjie Hu, Dan Garrette, Graham Neubig and Melvin Johnson	10215
<i>Contrastive Conditioning for Assessing Disambiguation in MT: A Case Study of Distilled Bias</i>	
Jannis Vamvas and Rico Sennrich	10246
<i>Measuring Association Between Labels and Free-Text Rationales</i>	
Sarah Wiegreffe, Ana Marasović and Noah A. Smith	10266
<i>Discretized Integrated Gradients for Explaining Language Models</i>	
Soumya Sanyal and Xiang Ren	10285
<i>Putting Words in BERT’s Mouth: Navigating Contextualized Vector Spaces with Pseudowords</i>	
Taelin Karidi, Yichu Zhou, Nathan Schneider, Omri Abend and Vivek Srikumar	10300
<i>Rationales for Sequential Predictions</i>	
Keyon Vafa, Yuntian Deng, David Blei and Alexander Rush	10314
<i>FastIF: Scalable Influence Functions for Efficient Model Interpretation and Debugging</i>	
Han Guo, Nazneen Rajani, Peter Hase, Mohit Bansal and Caiming Xiong	10333
<i>Studying word order through iterative shuffling</i>	
Nikolay Malkin, Sameera Lanka, Pranav Goel and Nebojsa Jojic	10351

<i>Distantly-Supervised Named Entity Recognition with Noise-Robust Learning and Language Model Augmented Self-Training</i>	
Yu Meng, Yunyi Zhang, Jiaxin Huang, Xuan Wang, Yu Zhang, Heng Ji and Jiawei Han	10367
<i>Open Knowledge Graphs Canonicalization using Variational Autoencoders</i>	
Sarthak Dash, Gaetano Rossiello, Nandana Mihindukulasooriya, Sugato Bagchi and Alfio Gliozzo	10379
<i>HittER: Hierarchical Transformers for Knowledge Graph Embeddings</i>	
Sanxing Chen, Xiaodong Liu, Jianfeng Gao, Jian Jiao, Ruofei Zhang and Yangfeng Ji	10395
<i>Few-Shot Named Entity Recognition: An Empirical Baseline Study</i>	
Jiaxin Huang, Chunyuan Li, Krishan Subudhi, Damien Jose, Shobana Balakrishnan, Weizhu Chen, Baolin Peng, Jianfeng Gao and Jiawei Han	10408
<i>XLEnt: Mining a Large Cross-lingual Entity Dataset with Lexical-Semantic-Phonetic Word Alignment</i>	
Ahmed El-Kishky, Adithya Renduchintala, James Cross, Francisco Guzmán and Philipp Koehn	10424
<i>Utilizing Relative Event Time to Enhance Event-Event Temporal Relation Extraction</i>	
Haoyang Wen and Heng Ji	10431
<i>Separating Retention from Extraction in the Evaluation of End-to-end Relation Extraction</i>	
Bruno Taillé, Vincent Guigue, Geoffrey Scoutheeten and patrick Gallinari	10438
<i>Automatic Text Evaluation through the Lens of Wasserstein Barycenters</i>	
Pierre Colombo, Guillaume Staerman, Chloé Clavel and Pablo Piantanida	10450
<i>Visually Grounded Reasoning across Languages and Cultures</i>	
Fangyu Liu, Emanuele Bugliarello, Edoardo Maria Ponti, Siva Reddy, Nigel Collier and Desmond Elliott	10467
<i>Back to Square One: Artifact Detection, Training and Commonsense Disentanglement in the Winograd Schema</i>	
Yanai Elazar, Hongming Zhang, Yoav Goldberg and Dan Roth	10486
<i>Robustness Evaluation of Entity Disambiguation Using Prior Probes: the Case of Entity Overshadowing</i>	
Vera Provatorova, Samarth Bhargav, Svitlana Vakulenko and Evangelos Kanoulas	10501
<i>IndoNLI: A Natural Language Inference Dataset for Indonesian</i>	
Rahmad Mahendra, Alham Fikri Aji, Samuel Louvan, Fahrurrozi Rahman and Clara Vania .	10511
<i>Agreeing to Disagree: Annotating Offensive Language Datasets with Annotators' Disagreement</i>	
Elisa Leonardelli, Stefano Menini, Alessio Palmero Aprosio, Marco Guerini and Sara Tonelli	10528
<i>A Root of a Problem: Optimizing Single-Root Dependency Parsing</i>	
Miloš Stanojević and Shay B. Cohen	10540
<i>Efficient Sampling of Dependency Structure</i>	
Ran Zmigrod, Tim Vieira and Ryan Cotterell	10558
<i>Reducing Discontinuous to Continuous Parsing with Pointer Network Reordering</i>	
Daniel Fernández-González and Carlos Gómez-Rodríguez	10570
<i>A New Representation for Span-based CCG Parsing</i>	
Yoshihide Kato and Shigeki Matsubara	10579

<i>What to Pre-Train on? Efficient Intermediate Task Selection</i>	
Clifton Poth, Jonas Pfeiffer, Andreas Rücklé and Iryna Gurevych	10585
<i>PermuteFormer: Efficient Relative Position Encoding for Long Sequences</i>	
Peng Chen	10606
<i>Block Pruning For Faster Transformers</i>	
François Lagunas, Ella Charlaix, Victor Sanh and Alexander Rush	10619
<i>Finetuning Pretrained Transformers into RNNs</i>	
Jungo Kasai, Hao Peng, Yizhe Zhang, Dani Yogatama, Gabriel Ilharco, Nikolaos Pappas, Yi Mao, Weizhu Chen and Noah A. Smith	10630
<i>How to Train BERT with an Academic Budget</i>	
Peter Izsak, Moshe Berchansky and Omer Levy	10644
<i>Beyond Preserved Accuracy: Evaluating Loyalty and Robustness of BERT Compression</i>	
Canwen Xu, Wangchunshu Zhou, Tao Ge, Ke Xu, Julian McAuley and Furu Wei	10653
<i>IndoBERTweet: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization</i>	
Fajri Koto, Jey Han Lau and Timothy Baldwin	10660
<i>Pushing on Text Readability Assessment: A Transformer Meets Handcrafted Linguistic Features</i>	
Bruce W. Lee, Yoo Sung Jang and Jason Lee	10669
<i>Types of Out-of-Distribution Texts and How to Detect Them</i>	
Udit Arora, William Huang and He He	10687
<i>Self-training with Few-shot Rationalization</i>	
Meghana Moorthy Bhat, Alessandro Sordoni and Subhabrata Mukherjee	10702
<i>MTAdam: Automatic Balancing of Multiple Training Loss Terms</i>	
Itzik Malkiel and Lior Wolf	10713
<i>Softmax Tree: An Accurate, Fast Classifier When the Number of Classes Is Large</i>	
Arman Zharmagambetov, Magzhan Gabidolla and Miguel A. Carreira-Perpinan	10730
<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	10746
<i>Multivalent Entailment Graphs for Question Answering</i>	
Nick McKenna, Liane Guillou, Mohammad Javad Hosseini, Sander Bijl de Vroe, Mark Johnson and Mark Steedman	10758
<i>Is Everything in Order? A Simple Way to Order Sentences</i>	
Somnath Basu Roy Chowdhury, Faeze Brahma and Snigdha Chaturvedi	10769
<i>VeeAlign: Multifaceted Context Representation Using Dual Attention for Ontology Alignment</i>	
Vivek Iyer, Arvind Agarwal and Harshit Kumar	10780
<i>Finding needles in a haystack: Sampling Structurally-diverse Training Sets from Synthetic Data for Compositional Generalization</i>	
Inbar Oren, Jonathan Herzig and Jonathan Berant	10793

<i>GeneSis: A Generative Approach to Substitutes in Context</i> Caterina Lacerra, Rocco Tripodi and Roberto Navigli	10810
<i>Semi-Supervised Exaggeration Detection of Health Science Press Releases</i> Dustin Wright and Isabelle Augenstein	10824
<i>Phrase-BERT: Improved Phrase Embeddings from BERT with an Application to Corpus Exploration</i> Shufan Wang, Laure Thompson and Mohit Iyyer	10837
<i>Detecting Contact-Induced Semantic Shifts: What Can Embedding-Based Methods Do in Practice?</i> Filip Miletic, Anne Przewozny-Desriaux and Ludovic Tanguy	10852