

16th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2021)

Held online due to COVID-19

Kiev, Ukraine
21 – 23 April 2021

Volume 1 of 5

ISBN: 978-1-7138-2914-0

Printed from e-media with permission by:

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



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

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

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

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

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

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

acl@aclweb.org

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>Unsupervised Sentence-embeddings by Manifold Approximation and Projection</i> Subhradeep Kayal	1
<i>Contrastive Multi-document Question Generation</i> Woon Sang Cho, Yizhe Zhang, Sudha Rao, Asli Celikyilmaz, Chenyan Xiong, Jianfeng Gao, Mengdi Wang and Bill Dolan	12
<i>Disambiguatory Signals are Stronger in Word-initial Positions</i> Tiago Pimentel, Ryan Cotterell and Brian Roark	31
<i>On the (In)Effectiveness of Images for Text Classification</i> Chunpeng Ma, Aili Shen, Hiyori Yoshikawa, Tomoya Iwakura, Daniel Beck and Timothy Baldwin 42	
<i>If you've got it, flaunt it: Making the most of fine-grained sentiment annotations</i> Jeremy Barnes, Lilja Øvrelid and Erik Velldal	49
<i>Keep Learning: Self-supervised Meta-learning for Learning from Inference</i> Akhil Kedia and SAI CHETAN CHINTHAKINDI	63
<i>ResPer: Computationally Modelling Resisting Strategies in Persuasive Conversations</i> Ritam Dutt, Sayan Sinha, Rishabh Joshi, Surya Shekhar Chakraborty, Meredith Riggs, Xinru Yan, Haogang Bao and Carolyn Rose	78
<i>BERxiT: Early Exiting for BERT with Better Fine-Tuning and Extension to Regression</i> Ji Xin, Raphael Tang, Yaoliang Yu and Jimmy Lin	91
<i>Telling BERT's Full Story: from Local Attention to Global Aggregation</i> Damian Pascual, Gino Brunner and Roger Wattenhofer	105
<i>Effects of Pre- and Post-Processing on type-based Embeddings in Lexical Semantic Change Detection</i> Jens Kaiser, Sinan Kurtiyigit, Serge Kotchourko and Dominik Schlechtweg	125
<i>The Gutenberg Dialogue Dataset</i> Richard Csaky and Gábor Recski	138
<i>On the Calibration and Uncertainty of Neural Learning to Rank Models for Conversational Search</i> Gustavo Penha and Claudia Hauff	160
<i>Frequency-Guided Word Substitutions for Detecting Textual Adversarial Examples</i> Maximilian Mozes, Pontus Stenetorp, Bennett Kleinberg and Lewis Griffin	171
<i>Maximal Multiverse Learning for Promoting Cross-Task Generalization of Fine-Tuned Language Models</i> Itzik Malkiel and Lior Wolf	187
<i>Unification-based Reconstruction of Multi-hop Explanations for Science Questions</i> Marco Valentino, Moganarangan Thayaparan and André Freitas	200
<i>Dictionary-based Debiasing of Pre-trained Word Embeddings</i> Masahiro Kaneko and Danushka Bollegala	212
<i>Belief-based Generation of Argumentative Claims</i> Milad Alshomary, Wei-Fan Chen, Timon Gurcke and Henning Wachsmuth	224

<i>Non-Autoregressive Text Generation with Pre-trained Language Models</i>	
Yixuan Su, Deng Cai, Yan Wang, David Vandyke, Simon Baker, Piji Li and Nigel Collier	234
<i>Multi-split Reversible Transformers Can Enhance Neural Machine Translation</i>	
Yuekai Zhao, Shuchang Zhou and Zhihua Zhang	244
<i>Exploiting Cloze-Questions for Few-Shot Text Classification and Natural Language Inference</i>	
Timo Schick and Hinrich Schütze	255
<i>CD²CR: Co-reference resolution across documents and domains</i>	
James Ravenscroft, Amanda Clare, Arie Cattan, Ido Dagan and Maria Liakata	270
<i>AREDSUM: Adaptive Redundancy-Aware Iterative Sentence Ranking for Extractive Document Summarization</i>	
Keping Bi, Rahul Jha, Bruce Croft and Asli Celikyilmaz	281
<i>“Talk to me with left, right, and angles”: Lexical entrainment in spoken Hebrew dialogue</i>	
Andreas Weise, Vered Silber-Varod, Anat Lerner, Julia Hirschberg and Rivka Levitan	292
<i>Recipes for Building an Open-Domain Chatbot</i>	
Stephen Roller, Emily Dinan, Naman Goyal, Da JU, Mary Williamson, Yinhan Liu, Jing Xu, Myle Ott, Eric Michael Smith, Y-Lan Boureau and Jason Weston	300
<i>Evaluating the Evaluation of Diversity in Natural Language Generation</i>	
Guy Tevet and Jonathan Berant	326
<i>Retrieval, Re-ranking and Multi-task Learning for Knowledge-Base Question Answering</i>	
Zhiguo Wang, Patrick Ng, Ramesh Nallapati and Bing Xiang	347
<i>Implicitly Abusive Comparisons – A New Dataset and Linguistic Analysis</i>	
Michael Wiegand, Maja Geulig and Josef Ruppenhofer	358
<i>Exploiting Emojis for Abusive Language Detection</i>	
Michael Wiegand and Josef Ruppenhofer	369
<i>A Systematic Review of Reproducibility Research in Natural Language Processing</i>	
Anya Belz, Shubham Agarwal, Anastasia Shimorina and Ehud Reiter	381
<i>Bootstrapping Multilingual AMR with Contextual Word Alignments</i>	
Janaki Sheth, Young-Suk Lee, Ramón Fernandez Astudillo, Tahira Naseem, Radu Florian, Salim Roukos and Todd Ward	394
<i>Semantic Oppositeness Assisted Deep Contextual Modeling for Automatic Rumor Detection in Social Networks</i>	
Nisansa de Silva and Dejing Dou	405
<i>Polarized-VAE: Proximity Based Disentangled Representation Learning for Text Generation</i>	
Vikash Balasubramanian, Ivan Kobyzev, Hareesh Bahuleyan, Ilya Shapiro and Olga Vechtomova	416
<i>ParaSCI: A Large Scientific Paraphrase Dataset for Longer Paraphrase Generation</i>	
Qingxiu Dong, Xiaojun Wan and Yue Cao	424
<i>Discourse Understanding and Factual Consistency in Abstractive Summarization</i>	
Saadia Gabriel, Antoine Bosselut, Jeff Da, Ari Holtzman, Jan Buys, Kyle Lo, Asli Celikyilmaz and Yejin Choi	435

<i>Knowledge Base Question Answering through Recursive Hypergraphs</i>	
Naganand yadati, Dayanidhi R S, Vaishnavi S, Indira K M and srinidhi g	448
<i>FEWS: Large-Scale, Low-Shot Word Sense Disambiguation with the Dictionary</i>	
Terra Blevins, Mandar Joshi and Luke Zettlemoyer	455
<i>MONAH: Multi-Modal Narratives for Humans to analyze conversations</i>	
Joshua Y. Kim, Kalina Yacef, Greyson Kim, Chunfeng Liu, Rafael Calvo and Silas Taylor	466
<i>Does Typological Blinding Impede Cross-Lingual Sharing?</i>	
Johannes Bjerva and Isabelle Augenstein	480
<i>AdapterFusion: Non-Destructive Task Composition for Transfer Learning</i>	
Jonas Pfeiffer, Aishwarya Kamath, Andreas Rücklé, Kyunghyun Cho and Iryna Gurevych	487
<i>CHOLAN: A Modular Approach for Neural Entity Linking on Wikipedia and Wikidata</i>	
Manoj Prabhakar Kannan Ravi, Kuldeep Singh, Isaiah Onando Mulang', Saeedeh Shekarpour, Johannes Hoffart and Jens Lehmann	504
<i>Grounding as a Collaborative Process</i>	
Luciana Benotti and Patrick Blackburn	515
<i>Does She Wink or Does She Nod? A Challenging Benchmark for Evaluating Word Understanding of Language Models</i>	
Lutfi Kerem Senel and Hinrich Schütze	532
<i>Joint Coreference Resolution and Character Linking for Multiparty Conversation</i>	
Jiaxin Bai, Hongming Zhang, Yangqiu Song and Kun Xu	539
<i>Improving Factual Consistency Between a Response and Persona Facts</i>	
Mohsen Mesgar, Edwin Simpson and Iryna Gurevych	549
<i>PolyLM: Learning about Polysemy through Language Modeling</i>	
Alan Ansell, Felipe Bravo-Marquez and Bernhard Pfahringer	563
<i>Predicting Treatment Outcome from Patient Texts: The Case of Internet-Based Cognitive Behavioural Therapy</i>	
Evangelia Gogoulou, Magnus Boman, Fehmi Ben Abdesslem, Nils Hentati Isacsson, Viktor Kaldo and Magnus Sahlgren	575
<i>Scalable Evaluation and Improvement of Document Set Expansion via Neural Positive-Unlabeled Learning</i>	
Alon Jacovi, Gang Niu, Yoav Goldberg and Masashi Sugiyama	581
<i>The Role of Syntactic Planning in Compositional Image Captioning</i>	
Emanuele Bugliarello and Desmond Elliott	593
<i>Is "hot pizza" Positive or Negative? Mining Target-aware Sentiment Lexicons</i>	
Jie Zhou, Yuanbin Wu, Changzhi Sun and Liang He	608
<i>Quality Estimation without Human-labeled Data</i>	
Yi-Lin Tuan, Ahmed El-Kishky, Adithya Renduchintala, Vishrav Chaudhary, Francisco Guzmán and Lucia Specia	619
<i>How Fast can BERT Learn Simple Natural Language Inference?</i>	
Yi-Chung Lin and Keh-Yih Su	626

<i>GRIT: Generative Role-filler Transformers for Document-level Event Entity Extraction</i>	
Xinya Du, Alexander Rush and Claire Cardie	634
<i>Cross-lingual Entity Alignment with Incidental Supervision</i>	
Muhao Chen, Weijia Shi, Ben Zhou and Dan Roth	645
<i>Query Generation for Multimodal Documents</i>	
kyungho kim, Kyungjae Lee, Seung-won Hwang, Young-In Song and seungwook lee	659
<i>End-to-End Argument Mining as Biaffine Dependency Parsing</i>	
Yuxiao Ye and Simone Teufel	669
<i>FakeFlow: Fake News Detection by Modeling the Flow of Affective Information</i>	
Bilal Ghanem, Simone Paolo Ponzetto, Paolo Rosso and Francisco Rangel	679
<i>CTC-based Compression for Direct Speech Translation</i>	
Marco Gaido, Mauro Cettolo, Matteo Negri and Marco Turchi	690
<i>A Crowdsourced Open-Source Kazakh Speech Corpus and Initial Speech Recognition Baseline</i>	
Yerbolat Khassanov, Saida Mussakhjayeva, Almas Mirzakhmetov, Alen Adiyev, Mukhamet Nurpei- issov and Huseyin Atakan Varol	697
<i>TDMSci: A Specialized Corpus for Scientific Literature Entity Tagging of Tasks Datasets and Metrics</i>	
Yufang Hou, Charles Jochim, Martin Gleize, Francesca Bonin and Debasis Ganguly	707
<i>Top-down Discourse Parsing via Sequence Labelling</i>	
Fajri Koto, Jey Han Lau and Timothy Baldwin	715
<i>Does the Order of Training Samples Matter? Improving Neural Data-to-Text Generation with Curriculum Learning</i>	
Ernie Chang, Hui-Syuan Yeh and Vera Demberg	727
<i>TrNews: Heterogeneous User-Interest Transfer Learning for News Recommendation</i>	
Guangneng Hu and Qiang Yang	734
<i>Dialogue Act-based Breakdown Detection in Negotiation Dialogues</i>	
Atsuki Yamaguchi, Kosui Iwasa and Katsuhide Fujita	745
<i>Neural Data-to-Text Generation with LM-based Text Augmentation</i>	
Ernie Chang, Xiaoyu Shen, Dawei Zhu, Vera Demberg and Hui Su	758
<i>Self-Training Pre-Trained Language Models for Zero- and Few-Shot Multi-Dialectal Arabic Sequence Labeling</i>	
Muhammad Khalifa, Muhammad Abdul-Mageed and Khaled Shaalan	769
<i>Multiple Tasks Integration: Tagging, Syntactic and Semantic Parsing as a Single Task</i>	
Timothée Bernard	783
<i>Coordinate Constructions in English Enhanced Universal Dependencies: Analysis and Computational Modeling</i>	
Stefan Grünewald, Prisca Piccirilli and Annemarie Friedrich	795
<i>Ellipsis Resolution as Question Answering: An Evaluation</i>	
Rahul Aralikkatte, Matthew Lamm, Daniel Hardt and Anders Søgaard	810

<i>Jointly Improving Language Understanding and Generation with Quality-Weighted Weak Supervision of Automatic Labeling</i>	
Ernie Chang, Vera Demberg and Alex Marin	818
<i>Continuous Learning in Neural Machine Translation using Bilingual Dictionaries</i>	
Jan Niehues	830
<i>Adv-OLM: Generating Textual Adversaries via OLM</i>	
Vijit Malik, Ashwani Bhat and Ashutosh Modi	841
<i>Conversational Question Answering over Knowledge Graphs with Transformer and Graph Attention Networks</i>	
Endri Kacupaj, Joan Plepi, Kuldeep Singh, Harsh Thakkar, Jens Lehmann and Maria Maleshkova	850
<i>DRAG: Director-Generator Language Modelling Framework for Non-Parallel Author Stylized Rewriting</i>	
Hrituraj Singh, Gaurav Verma, Aparna Garimella and Balaji Vasan Srinivasan	863
<i>Leveraging Passage Retrieval with Generative Models for Open Domain Question Answering</i>	
Gautier Izacard and Edouard Grave	874
<i>Clinical Outcome Prediction from Admission Notes using Self-Supervised Knowledge Integration</i>	
Betty van Aken, Jens-Michalis Papaioannou, Manuel Mayrdorfer, Klemens Budde, Felix Gers and Alexander Loeser	881
<i>Combining Deep Generative Models and Multi-lingual Pretraining for Semi-supervised Document Classification</i>	
Yi Zhu, Ehsan Shareghi, Yingzhen Li, Roi Reichart and Anna Korhonen	894
<i>Multi-facet Universal Schema</i>	
Rohan Paul, Haw-Shiuan Chang and Andrew McCallum	909
<i>Exploring Transitivity in Neural NLI Models through Veridicality</i>	
Hitomi Yanaka, Koji Mineshima and Kentaro Inui	920
<i>A Neural Few-Shot Text Classification Reality Check</i>	
Thomas Dopierre, Christophe Gravier and Wilfried Logerais	935
<i>Multilingual Machine Translation: Closing the Gap between Shared and Language-specific Encoder-Decoders</i>	
Carlos Escolano, Marta R. Costa-jussà, José A. R. Fonollosa and Mikel Artetxe	944
<i>Clustering Word Embeddings with Self-Organizing Maps. Application on LaRoSeDa - A Large Romanian Sentiment Data Set</i>	
Anca Tache, Gaman Mihaela and Radu Tudor Ionescu	949
<i>Elastic weight consolidation for better bias inoculation</i>	
James Thorne and Andreas Vlachos	957
<i>Hierarchical Multi-head Attentive Network for Evidence-aware Fake News Detection</i>	
Nguyen Vo and Kyumin Lee	965
<i>Identifying Named Entities as they are Typed</i>	
Ravneet Arora, Chen-Tse Tsai and Daniel Preotiuc-Pietro	976

<i>SANDI: Story-and-Images Alignment</i>	
Sreyasi Nag Chowdhury, Simon Razniewski and Gerhard Weikum	989
<i>Question and Answer Test-Train Overlap in Open-Domain Question Answering Datasets</i>	
Patrick Lewis, Pontus Stenetorp and Sebastian Riedel	1000
<i>El Volumen Louder Por Favor: Code-switching in Task-oriented Semantic Parsing</i>	
Arash Einolghozati, Abhinav Arora, Lorena Sainz-Maza Lecanda, Anuj Kumar and Sonal Gupta	1009
<i>Generating Syntactically Controlled Paraphrases without Using Annotated Parallel Pairs</i>	
Kuan-Hao Huang and Kai-Wei Chang	1022
<i>Data Augmentation for Hypernymy Detection</i>	
Thomas Kober, Julie Weeds, Lorenzo Bertolini and David Weir	1034
<i>Few-shot learning through contextual data augmentation</i>	
Farid Arthaud, Rachel Bawden and Alexandra Birch	1049
<i>Zero-shot Generalization in Dialog State Tracking through Generative Question Answering</i>	
Shuyang Li, Jin Cao, Mukund Sridhar, Henghui Zhu, Shang-Wen Li, Wael Hamza and Julian McAuley	1063
<i>Zero-shot Neural Passage Retrieval via Domain-targeted Synthetic Question Generation</i>	
Ji Ma, Ivan Korotkov, Yinfei Yang, Keith Hall and Ryan McDonald	1075
<i>Discourse-Aware Unsupervised Summarization for Long Scientific Documents</i>	
Yue Dong, Andrei Mircea Romascanu and Jackie Chi Kit Cheung	1089
<i>MIDAS: A Dialog Act Annotation Scheme for Open Domain HumanMachine Spoken Conversations</i>	
Dian Yu and Zhou Yu	1103
<i>Analyzing the Forgetting Problem in Pretrain-Finetuning of Open-domain Dialogue Response Models</i>	
Tianxing He, Jun Liu, Kyunghyun Cho, Myle Ott, Bing Liu, James Glass and Fuchun Peng . .	1121
<i>Leveraging End-to-End ASR for Endangered Language Documentation: An Empirical Study on Yolóxo-chitl Mixtec</i>	
Jiatong Shi, Jonathan D. Amith, Rey Castillo García, Esteban Guadalupe Sierra, Kevin Duh and Shinji Watanabe	1134
<i>Mode Effects' Challenge to Authorship Attribution</i>	
Haining Wang, Allen Riddell and Patrick Juola	1146
<i>Generative Text Modeling through Short Run Inference</i>	
Bo Pang, Erik Nijkamp, Tian Han and Ying Nian Wu	1156
<i>Detecting Extraneous Content in Podcasts</i>	
Sravana Reddy, Yongze Yu, Aasish Pappu, Aswin Sivaraman, Rezvaneh Rezapour and Rosie Jones	1166
<i>Randomized Deep Structured Prediction for Discourse-Level Processing</i>	
Manuel Widmoser, Maria Pacheco, Jean Honorio and Dan Goldwasser	1174
<i>Automatic Data Acquisition for Event Coreference Resolution</i>	
Prafulla Kumar Choubey and Ruihong Huang	1185

<i>Joint Learning of Representations for Web-tables, Entities and Types using Graph Convolutional Network</i>	
Aniket Pramanick and Indrajit Bhattacharya	1197
<i>Multimodal Text Style Transfer for Outdoor Vision-and-Language Navigation</i>	
Wanrong Zhu, Xin Wang, Tsu-Jui Fu, An Yan, Pradyumna Narayana, Kazoo Sone, Sugato Basu and William Yang Wang	1207
<i>ECOL-R: Encouraging Copying in Novel Object Captioning with Reinforcement Learning</i>	
Yufei Wang, Ian Wood, Stephen Wan and Mark Johnson	1222
<i>Enriching Non-Autoregressive Transformer with Syntactic and Semantic Structures for Neural Machine Translation</i>	
Ye Liu, Yao Wan, Jianguo Zhang, Wenting Zhao and Philip Yu	1235
<i>NLQuAD: A Non-Factoid Long Question Answering Data Set</i>	
Amir Soleimani, Christof Monz and marcel worring	1245
<i>Debiasing Pre-trained Contextualised Embeddings</i>	
Masahiro Kaneko and Danushka Bollegala	1256
<i>Language Models for Lexical Inference in Context</i>	
Martin Schmitt and Hinrich Schütze	1267
<i>Few-Shot Semantic Parsing for New Predicates</i>	
Zhuang Li, Lizhen Qu, shuo huang and Gholamreza Haffari	1281
<i>Alternating Recurrent Dialog Model with Large-scale Pre-trained Language Models</i>	
Qingyang Wu, Yichi Zhang, Yu Li and Zhou Yu	1292
<i>On the Evaluation of Vision-and-Language Navigation Instructions</i>	
Ming Zhao, Peter Anderson, Vihan Jain, Su Wang, Alexander Ku, Jason Baldridge and Eugene Ie	1302
<i>Cross-lingual Visual Pre-training for Multimodal Machine Translation</i>	
Ozan Caglayan, Menekse Kuyu, Mustafa Serkan Amac, Pranava Madhyastha, Erkut Erdem, Aykut Erdem and Lucia Specia	1317
<i>Memorization vs. Generalization : Quantifying Data Leakage in NLP Performance Evaluation</i>	
Aparna Elangovan, Jiayuan He and Karin Verspoor	1325
<i>An Expert Annotated Dataset for the Detection of Online Misogyny</i>	
Ella Guest, Bertie Vidgen, Alexandros Mittos, Nishanth Sastry, Gareth Tyson and Helen Margetts	1336
<i>WikiMatrix: Mining 135M Parallel Sentences in 1620 Language Pairs from Wikipedia</i>	
Holger Schwenk, Vishrav Chaudhary, Shuo Sun, Hongyu Gong and Francisco Guzmán	1351
<i>ChEMU-Ref: A Corpus for Modeling Anaphora Resolution in the Chemical Domain</i>	
Biaoyan Fang, Christian Druckenbrodt, Saber A Akhondi, Jiayuan He, Timothy Baldwin and Karin Verspoor	1362
<i>Syntactic Nuclei in Dependency Parsing – A Multilingual Exploration</i>	
Ali Basirat and Joakim Nivre	1376

<i>Searching for Search Errors in Neural Morphological Inflection</i>	
Martina Forster, Clara Meister and Ryan Cotterell	1388
<i>Quantifying Appropriateness of Summarization Data for Curriculum Learning</i>	
Ryuji Kano, Takumi Takahashi, Toru Nishino, Motoki Taniguchi, Tomoki Taniguchi and Tomoko Ohkuma.....	1395
<i>Evaluating language models for the retrieval and categorization of lexical collocations</i>	
Luis Espinosa Anke, Joan Codina-Filba and Leo Wanner	1406
<i>BART-TL: Weakly-Supervised Topic Label Generation</i>	
Cristian Popa and Traian Rebedea	1418
<i>Dynamic Graph Transformer for Implicit Tag Recognition</i>	
Yi-Ting Liou, Chung-Chi Chen, Hen-Hsen Huang and Hsin-Hsi Chen	1426
<i>Implicit Unlikelihood Training: Improving Neural Text Generation with Reinforcement Learning</i>	
Evgeny Lagutin, Daniil Gavrilov and Pavel Kalaidin	1432
<i>Civil Rephrases Of Toxic Texts With Self-Supervised Transformers</i>	
Léo Laugier, John Pavlopoulos, Jeffrey Sorensen and Lucas Dixon	1442
<i>Generating Weather Comments from Meteorological Simulations</i>	
Soichiro Murakami, Sora Tanaka, Masatsugu Hangyo, Hidetaka Kamigaito, Kotaro Funakoshi, Hiroya Takamura and Manabu Okumura.....	1462
<i>SICK-NL: A Dataset for Dutch Natural Language Inference</i>	
Gijs Wijnholds and Michael Moortgat	1474
<i>A phonetic model of non-native spoken word processing</i>	
Yevgen Matushevych, Herman Kamper, Thomas Schatz, Naomi Feldman and Sharon Goldwater	1480
<i>Bootstrapping Relation Extractors using Syntactic Search by Examples</i>	
Matan Eyal, Asaf Amrami, Hillel Taub-Tabib and Yoav Goldberg	1491
<i>Towards a Decomposable Metric for Explainable Evaluation of Text Generation from AMR</i>	
Juri Opitz and Anette Frank	1504
<i>The Source-Target Domain Mismatch Problem in Machine Translation</i>	
Jiajun Shen, Peng-Jen Chen, Matthew Le, Junxian He, Jiatao Gu, Myle Ott, Michael Auli and Marc’Aurelio Ranzato	1519
<i>Cross-Topic Rumor Detection using Topic-Mixtures</i>	
Xiaoying Ren, Jing Jiang, Ling Min Serena Khoo and Hai Leong Chieu.....	1534
<i>Understanding Pre-Editing for Black-Box Neural Machine Translation</i>	
Rei Miyata and Atsushi Fujita	1539
<i>RelWalk - A Latent Variable Model Approach to Knowledge Graph Embedding</i>	
Danushka Bollegala, Huda Hakami, Yuichi Yoshida and Ken-ichi Kawarabayashi.....	1551
<i>Few-shot Learning for Slot Tagging with Attentive Relational Network</i>	
Cennet Oguz and Ngoc Thang Vu	1566

<i>SpanEmo: Casting Multi-label Emotion Classification as Span-prediction</i>	
Hassan Alhuzali and Sophia Ananiadou	1573
<i>Exploiting Position and Contextual Word Embeddings for Keyphrase Extraction from Scientific Papers</i>	
Krutarth Patel and Cornelia Caragea	1585
<i>Benchmarking Machine Reading Comprehension: A Psychological Perspective</i>	
Saku Sugawara, Pontus Stenetorp and Akiko Aizawa	1592
<i>Multilingual Neural Machine Translation with Deep Encoder and Multiple Shallow Decoders</i>	
Xiang Kong, Adithya Renduchintala, James Cross, Yuqing Tang, Jiatao Gu and Xian Li	1613
<i>With Measured Words: Simple Sentence Selection for Black-Box Optimization of Sentence Compression Algorithms</i>	
Yotam Shichel, Meir Kalech and Oren Tsur	1625
<i>WiC-TSV: An Evaluation Benchmark for Target Sense Verification of Words in Context</i>	
Anna Breit, Artem Revenko, Kiamehr Rezaee, Mohammad Taher Pilehvar and Jose Camacho-Collados	1635
<i>Self-Supervised and Controlled Multi-Document Opinion Summarization</i>	
Hady Elsahar, Maximin Coavoux, Jos Rozen and Matthias Gallé	1646
<i>NewsMTSC: A Dataset for (Multi-)Target-dependent Sentiment Classification in Political News Articles</i>	
Felix Hamborg and Karsten Donnay	1663
<i>Cross-lingual Contextualized Topic Models with Zero-shot Learning</i>	
Federico Bianchi, Silvia Terragni, Dirk Hovy, Debora Nozza and Elisabetta Fersini	1676
<i>Dependency parsing with structure preserving embeddings</i>	
Ákos Kádár, Lan Xiao, Mete Kemertas, Federico Fancellu, Allan Jepson and Afsaneh Fazly	1684
<i>Active Learning for Sequence Tagging with Deep Pre-trained Models and Bayesian Uncertainty Estimates</i>	
Artem Shelmanov, Dmitri Puzyrev, Lyubov Kupriyanova, Denis Belyakov, Daniil Larionov, Nikita Khromov, Olga Kozlova, Ekaterina Artemova, Dmitry V. Dylov and Alexander Panchenko	1698
<i>MultiHumES: Multilingual Humanitarian Dataset for Extractive Summarization</i>	
Jenny Paola Yela-Bello, Ewan Oglethorpe and Navid Rekabsaz	1713
<i>Learning From Revisions: Quality Assessment of Claims in Argumentation at Scale</i>	
Gabriella Skitalinskaya, Jonas Klaff and Henning Wachsmuth	1718
<i>Few Shot Dialogue State Tracking using Meta-learning</i>	
Saket Dingliwal, Shuyang Gao, Sanchit Agarwal, Chien-Wei Lin, Tagyoung Chung and Dilek Hakkani-Tur	1730
<i>BERT Prescriptions to Avoid Unwanted Headaches: A Comparison of Transformer Architectures for Adverse Drug Event Detection</i>	
Beatrice Portelli, Edoardo Lenzi, Emmanuele Chersoni, Giuseppe Serra and Enrico Santus	1740
<i>Semantic Parsing of Disfluent Speech</i>	
Priyanka Sen and Isabel Groves	1748
<i>Joint Energy-based Model Training for Better Calibrated Natural Language Understanding Models</i>	
Tianxing He, Bryan McCann, Caiming Xiong and Ehsan Hosseini-Asl	1754

<i>What Sounds “Right” to Me? Experiential Factors in the Perception of Political Ideology</i>	
Qinlan Shen and Carolyn Rose	1762
<i>Language Models as Knowledge Bases: On Entity Representations, Storage Capacity, and Paraphrased Queries</i>	
Benjamin Heinzerling and Kentaro Inui	1772
<i>Globalizing BERT-based Transformer Architectures for Long Document Summarization</i>	
quentin grail, Julien PEREZ and Eric Gaussier	1792
<i>Through the Looking Glass: Learning to Attribute Synthetic Text Generated by Language Models</i>	
Shaoor Munir, Brishna Batool, Zubair Shafiq, Padmini Srinivasan and Fareed Zaffar	1811
<i>We Need To Talk About Random Splits</i>	
Anders Søgaard, Sebastian Ebert, Jasmijn Bastings and Katja Filippova	1823
<i>How Certain is Your Transformer?</i>	
Artem Shelmanov, Evgenii Tsymbalov, Dmitri Puzyrev, Kirill Fedyanin, Alexander Panchenko and Maxim Panov	1833
<i>Alignment verification to improve NMT translation towards highly inflectional languages with limited resources</i>	
George Tambouratzis	1841
<i>Data Augmentation for Voice-Assistant NLU using BERT-based Interchangeable Rephrase</i>	
Akhila Yerukola, Mason Bretan and Hongxia Jin	1852
<i>How to Evaluate a Summarizer: Study Design and Statistical Analysis for Manual Linguistic Quality Evaluation</i>	
Julius Steen and Katja Markert	1861
<i>Open-Mindedness and Style Coordination in Argumentative Discussions</i>	
Aviv Ben-Haim and Oren Tsur	1876
<i>Error Analysis and the Role of Morphology</i>	
Marcel Bollmann and Anders Søgaard	1887
<i>Applying the Transformer to Character-level Transduction</i>	
Shijie Wu, Ryan Cotterell and Mans Hulden	1901
<i>Exploring Supervised and Unsupervised Rewards in Machine Translation</i>	
Julia Ive, Zixu Wang, Marina Fomicheva and Lucia Specia	1908
<i>Us vs. Them: A Dataset of Populist Attitudes, News Bias and Emotions</i>	
Pere-Lluís Huguet Cabot, David Abadi, Agneta Fischer and Ekaterina Shutova	1921
<i>Multilingual Entity and Relation Extraction Dataset and Model</i>	
Alessandro Seganti, Klaudia Firląg, Helena Skowronska, Michał Satława and Piotr Andruszkiewicz	
1946	
<i>A New View of Multi-modal Language Analysis: Audio and Video Features as Text “Styles”</i>	
Zhongkai Sun, Prathusha K Sarma, Yingyu Liang and William Sethares	1956
<i>Multilingual and cross-lingual document classification: A meta-learning approach</i>	
Niels van der Heijden, Helen Yannakoudakis, Pushkar Mishra and Ekaterina Shutova	1966

<i>Boosting Low-Resource Biomedical QA via Entity-Aware Masking Strategies</i>	
Gabriele Pergola, Elena Kochkina, Lin Gui, Maria Liakata and Yulan He	1977
<i>Attention-based Relational Graph Convolutional Network for Target-Oriented Opinion Words Extraction</i>	
Junfeng Jiang, An Wang and Akiko Aizawa	1986
<i>“Laughing at you or with you”: The Role of Sarcasm in Shaping the Disagreement Space</i>	
Debanjan Ghosh, Ritvik Shrivastava and Smaranda Muresan	1998
<i>Learning Relatedness between Types with Prototypes for Relation Extraction</i>	
Lisheng Fu and Ralph Grishman	2011
<i>I Beg to Differ: A study of constructive disagreement in online conversations</i>	
Christine De Kock and Andreas Vlachos	2017
<i>Acquiring a Formality-Informed Lexical Resource for Style Analysis</i>	
Elisabeth Eder, Ulrike Krieg-Holz and Udo Hahn	2028
<i>Probing into the Root: A Dataset for Reason Extraction of Structural Events from Financial Documents</i>	
Pei Chen, Kang Liu, Yubo Chen, Taifeng Wang and Jun Zhao	2042
<i>Language Modelling as a Multi-Task Problem</i>	
Lucas Weber, Jaap Jumelet, Elia Bruni and Dieuwke Hupkes	2049
<i>ChainCQG: Flow-Aware Conversational Question Generation</i>	
Jing Gu, Mostafa Mirshekari, Zhou Yu and Aaron Sisto	2061
<i>The Interplay of Task Success and Dialogue Quality: An in-depth Evaluation in Task-Oriented Visual Dialogues</i>	
Alberto Testoni and Raffaella Bernardi	2071
<i>"Are you kidding me?": Detecting Unpalatable Questions on Reddit</i>	
Sunyam Bagga, Andrew Piper and Derek Ruths	2083
<i>Neural-Driven Search-Based Paraphrase Generation</i>	
Betty Fabre, Tanguy Urvoy, Jonathan Chevelu and Damien Lolive	2100
<i>Word Alignment by Fine-tuning Embeddings on Parallel Corpora</i>	
Zi-Yi Dou and Graham Neubig	2112
<i>Paraphrases do not explain word analogies</i>	
Louis Fournier and Ewan Dunbar	2129
<i>An Empirical Study on the Generalization Power of Neural Representations Learned via Visual Guessing Games</i>	
Alessandro Suglia, Yonatan Bisk, Ioannis Konstas, Antonio Vergari, Emanuele Bastianelli, Andrea Vanzo and Oliver Lemon	2135
<i>A Unified Feature Representation for Lexical Connotations</i>	
Emily Allaway and Kathleen McKeown	2145
<i>FAST: Financial News and Tweet Based Time Aware Network for Stock Trading</i>	
Ramit Sawhney, Arnav Wadhwa, Shivam Agarwal and Rajiv Ratn Shah	2164

<i>Building Representative Corpora from Illiterate Communities: A Review of Challenges and Mitigation Strategies for Developing Countries</i>	
Stephanie Hirmer, Alycia Leonard, Josephine Tumwesige and Costanza Conforti	2176
<i>Process-Level Representation of Scientific Protocols with Interactive Annotation</i>	
Ronen Tamari, Fan Bai, Alan Ritter and Gabriel Stanovsky	2190
<i>Machine Translationese: Effects of Algorithmic Bias on Linguistic Complexity in Machine Translation</i>	
Eva Vanmassenhove, Dimitar Shterionov and Matthew Gwilliam	2203
<i>First Align, then Predict: Understanding the Cross-Lingual Ability of Multilingual BERT</i>	
Benjamin Muller, Yanai Elazar, Benoît Sagot and Djamé Seddah	2214
<i>Stereotype and Skew: Quantifying Gender Bias in Pre-trained and Fine-tuned Language Models</i>	
Daniel de Vassimon Manela, David Errington, Thomas Fisher, Boris van Breugel and Pasquale Minervini	2232
<i>On the evolution of syntactic information encoded by BERT's contextualized representations</i>	
Laura Pérez-Mayos, Roberto Carlini, Miguel Ballesteros and Leo Wanner	2243
<i>Identify, Align, and Integrate: Matching Knowledge Graphs to Commonsense Reasoning Tasks</i>	
Lisa Bauer and Mohit Bansal	2259
<i>Calculating the optimal step of arc-eager parsing for non-projective trees</i>	
Mark-Jan Nederhof	2273
<i>Subword Pooling Makes a Difference</i>	
Judit Ács, Ákos Kádár and Andras Kornai	2284
<i>Content-based Models of Quotation</i>	
Ansel MacLaughlin and David Smith	2296
<i>L2C: Describing Visual Differences Needs Semantic Understanding of Individuals</i>	
An Yan, Xin Wang, Tsu-Jui Fu and William Yang Wang	2315
<i>VoiSeR: A New Benchmark for Voice-Based Search Refinement</i>	
Simone Filice, Giuseppe Castellucci, Marcus Collins, Eugene Agichtein and Oleg Rokhlenko	2321
<i>Event-Driven News Stream Clustering using Entity-Aware Contextual Embeddings</i>	
Kailash Karthik Saravanakumar, Miguel Ballesteros, Muthu Kumar Chandrasekaran and Kathleen McKeown	2330
<i>Adversarial Learning of Poisson Factorisation Model for Gauging Brand Sentiment in User Reviews</i>	
Runcong Zhao, Lin Gui, Gabriele Pergola and Yulan He	2341
<i>Lexical Normalization for Code-switched Data and its Effect on POS Tagging</i>	
Rob van der Goot and Özlem Çetinoğlu	2352
<i>Structural Encoding and Pre-training Matter: Adapting BERT for Table-Based Fact Verification</i>	
Rui Dong and David Smith	2366
<i>A Study of Automatic Metrics for the Evaluation of Natural Language Explanations</i>	
Miruna-Adriana Clinciu, Arash Eshghi and Helen Hastie	2376
<i>Adversarial Stylometry in the Wild: Transferable Lexical Substitution Attacks on Author Profiling</i>	
Chris Emmery, Ákos Kádár and Grzegorz Chrupała	2388

<i>Cross-Cultural Similarity Features for Cross-Lingual Transfer Learning of Pragmatically Motivated Tasks</i>	
Jimin Sun, Hwijeen Ahn, Chan Young Park, Yulia Tsvetkov and David R. Mortensen	2403
<i>PHASE: Learning Emotional Phase-aware Representations for Suicide Ideation Detection on Social Media</i>	
Ramit Sawhney, Harshit Joshi, Lucie Flek and Rajiv Ratn Shah	2415
<i>Exploiting Definitions for Frame Identification</i>	
Tianyu Jiang and Ellen Riloff	2429
<i>ADePT: Auto-encoder based Differentially Private Text Transformation</i>	
Satyapriya Krishna, Rahul Gupta and Christophe Dupuy	2435
<i>Conceptual Grounding Constraints for Truly Robust Biomedical Name Representations</i>	
Pieter Fivez, Simon Suster and Walter Daelemans	2440
<i>Adaptive Mixed Component LDA for Low Resource Topic Modeling</i>	
Suzanna Sia and Kevin Duh	2451
<i>Evaluating Neural Model Robustness for Machine Comprehension</i>	
Winston Wu, Dustin Arendt and Svitlana Volkova	2470
<i>Hidden Biases in Unreliable News Detection Datasets</i>	
Xiang Zhou, Heba Elfardy, Christos Christodoulopoulos, Thomas Butler and Mohit Bansal . .	2482
<i>Annealing Knowledge Distillation</i>	
Aref Jafari, Mehdi Rezagholizadeh, Pranav Sharma and Ali Ghodsi	2493
<i>Unsupervised Extractive Summarization using Pointwise Mutual Information</i>	
Vishakh Padmakumar and He He	2505
<i>Context-aware Neural Machine Translation with Mini-batch Embedding</i>	
Makoto Morishita, Jun Suzuki, Tomoharu Iwata and Masaaki Nagata	2513
<i>Deep Subjecthood: Higher-Order Grammatical Features in Multilingual BERT</i>	
Isabel Papadimitriou, Ethan A. Chi, Richard Futrell and Kyle Mahowald	2522
<i>Streaming Models for Joint Speech Recognition and Translation</i>	
Orion Weller, Matthias Sperber, Christian Gollan and Joris Kluivers	2533
<i>DOCENT: Learning Self-Supervised Entity Representations from Large Document Collections</i>	
Yury Zemlyanskiy, Sudeep Gandhe, Ruining He, Bhargav Kanagal, Anirudh Ravula, Juraj Got-tweis, Fei Sha and Ilya Eckstein	2540
<i>Scientific Discourse Tagging for Evidence Extraction</i>	
Xiangci Li, Gully Burns and Nanyun Peng	2550
<i>Incremental Beam Manipulation for Natural Language Generation</i>	
James Hargreaves, Andreas Vlachos and Guy Emerson	2563
<i>StructSum: Summarization via Structured Representations</i>	
Vidhisha Balachandran, Artidoro Pagnoni, Jay Yoon Lee, Dheeraj Rajagopal, Jaime Carbonell and Yulia Tsvetkov	2575

<i>Project-then-Transfer: Effective Two-stage Cross-lingual Transfer for Semantic Dependency Parsing</i> Hiroaki Ozaki, Gaku Morio, Terufumi Morishita and Toshinori Miyoshi	2586
<i>LSOIE: A Large-Scale Dataset for Supervised Open Information Extraction</i> Jacob Solawetz and Stefan Larson	2595
<i>Changing the Mind of Transformers for Topically-Controllable Language Generation</i> Haw-Shiuan Chang, Jiaming Yuan, Mohit Iyyer and Andrew McCallum	2601
<i>Unsupervised Abstractive Summarization of Bengali Text Documents</i> Radia Rayan Chowdhury, Mir Tafseer Nayeem, Tahsin Tasnim Mim, Md. Saifur Rahman Chowdhury and Taufiqul Jannat	2612
<i>From Toxicity in Online Comments to Incivility in American News: Proceed with Caution</i> Anushree Hede, Oshin Agarwal, Linda Lu, Diana C. Mutz and Ani Nenkova	2620
<i>On the Computational Modelling of Michif Verbal Morphology</i> Fineen Davis, Eddie Antonio Santos and Heather Souter	2631
<i>A Few Topical Tweets are Enough for Effective User Stance Detection</i> Younes Samih and Kareem Darwish	2637
<i>Do Syntax Trees Help Pre-trained Transformers Extract Information?</i> Devendra Sachan, Yuhao Zhang, Peng Qi and William L. Hamilton	2647
<i>Informative and Controllable Opinion Summarization</i> Reinald Kim Amplayo and Mirella Lapata	2662
<i>Coloring the Black Box: What Synesthesia Tells Us about Character Embeddings</i> Katharina Kann and Mauro M. Monsalve-Mercado	2673
<i>How Good (really) are Grammatical Error Correction Systems?</i> Alla Rozovskaya and Dan Roth	2686
<i>BERTective: Language Models and Contextual Information for Deception Detection</i> Tommaso Fornaciari, Federico Bianchi, Massimo Poesio and Dirk Hovy	2699
<i>Learning Coupled Policies for Simultaneous Machine Translation using Imitation Learning</i> Philip Arthur, Trevor Cohn and Gholamreza Haffari	2709
<i>Complementary Evidence Identification in Open-Domain Question Answering</i> Xiangyang Mou, Mo Yu, Shiyu Chang, Yufei Feng, Li Zhang and Hui Su	2720
<i>Entity-level Factual Consistency of Abstractive Text Summarization</i> Feng Nan, Ramesh Nallapati, Zhiguo Wang, Cicero Nogueira dos Santos, Henghui Zhu, Dejiao Zhang, Kathleen McKeown and Bing Xiang	2727
<i>On Hallucination and Predictive Uncertainty in Conditional Language Generation</i> Yijun Xiao and William Yang Wang	2734
<i>Fine-Grained Event Trigger Detection</i> Duong Le and Thien Huu Nguyen	2745
<i>Extremely Small BERT Models from Mixed-Vocabulary Training</i> Sanqiang Zhao, Raghav Gupta, Yang Song and Denny Zhou	2753

<i>Diverse Adversaries for Mitigating Bias in Training</i>	
Xudong Han, Timothy Baldwin and Trevor Cohn	2760
<i>‘Just because you are right, doesn’t mean I am wrong’: Overcoming a bottleneck in development and evaluation of Open-Ended VQA tasks</i>	
Man Luo, Shailaja Keyur Sampat, Riley Tallman, Yankai Zeng, Manuha Vancha, Akarshan Sajja and Chitta Baral	2766
<i>Better Neural Machine Translation by Extracting Linguistic Information from BERT</i>	
Hassan S. Shavarani and Anoop Sarkar	2772
<i>CLiMP: A Benchmark for Chinese Language Model Evaluation</i>	
Beilei Xiang, Changbing Yang, Yu Li, Alex Warstadt and Katharina Kann	2784
<i>Measuring and Improving Faithfulness of Attention in Neural Machine Translation</i>	
Pooya Moradi, Nishant Kambhatla and Anoop Sarkar	2791
<i>Progressively Pretrained Dense Corpus Index for Open-Domain Question Answering</i>	
Wenhan Xiong, Hong Wang and William Yang Wang	2803
<i>Exploring the Limits of Few-Shot Link Prediction in Knowledge Graphs</i>	
Dora Jambor, Komal Teru, Joelle Pineau and William L. Hamilton	2816
<i>ProFormer: Towards On-Device LSH Projection Based Transformers</i>	
Chinnadhurai Sankar, Sujith Ravi and Zornitsa Kozareva	2823
<i>Joint Learning of Hyperbolic Label Embeddings for Hierarchical Multi-label Classification</i>	
Soumya Chatterjee, Ayush Maheshwari, Ganesh Ramakrishnan and Saketha Nath Jagaralpudi	2829
<i>Segmenting Subtitles for Correcting ASR Segmentation Errors</i>	
David Wan, Chris Kedzie, Faisal Ladhak, Elsbeth Turcan, Petra Galuscakova, Elena Zotkina, Zhengping Jiang, Peter Bell and Kathleen McKeown	2842
<i>Crisscrossed Captions: Extended Intramodal and Intermodal Semantic Similarity Judgments for MS-COCO</i>	
Zarana Parekh, Jason Baldridge, Daniel Cer, Austin Waters and Yinfei Yang	2855
<i>On-Device Text Representations Robust To Misspellings via Projections</i>	
Chinnadhurai Sankar, Sujith Ravi and Zornitsa Kozareva	2871
<i>ENPAR: Enhancing Entity and Entity Pair Representations for Joint Entity Relation Extraction</i>	
Yijun Wang, Changzhi Sun, Yuanbin Wu, Hao Zhou, Lei Li and Junchi Yan	2877
<i>Text Augmentation in a Multi-Task View</i>	
Jason Wei, Chengyu Huang, Shiqi Xu and Soroush Vosoughi	2888
<i>Representations for Question Answering from Documents with Tables and Text</i>	
Vicky Zayats, Kristina Toutanova and Mari Ostendorf	2895
<i>PPT: Parsimonious Parser Transfer for Unsupervised Cross-Lingual Adaptation</i>	
Kemal Kurniawan, Lea Frermann, Philip Schulz and Trevor Cohn	2907
<i>Modelling Context Emotions using Multi-task Learning for Emotion Controlled Dialog Generation</i>	
Deeksha Varshney, Asif Ekbal and Pushpak Bhattacharyya	2919

<i>Gender and Racial Fairness in Depression Research using Social Media</i>	
Carlos Aguirre, Keith Harrigian and Mark Dredze	2932
<i>MTOP: A Comprehensive Multilingual Task-Oriented Semantic Parsing Benchmark</i>	
Haoran Li, Abhinav Arora, Shuohui Chen, Anchit Gupta, Sonal Gupta and Yashar Mehdad ..	2950
<i>Adapting Event Extractors to Medical Data: Bridging the Covariate Shift</i>	
Aakanksha Naik, Jill Fain Lehman and Carolyn Rose	2963
<i>NoiseQA: Challenge Set Evaluation for User-Centric Question Answering</i>	
Abhilasha Ravichander, Siddharth Dalmia, Maria Ryskina, Florian Metze, Eduard Hovy and Alan W Black	2976
<i>Co-evolution of language and agents in referential games</i>	
Gautier Dagan, Dieuwke Hupkes and Elia Bruni	2993
<i>Modeling Context in Answer Sentence Selection Systems on a Latency Budget</i>	
Rujun Han, Luca Soldaini and Alessandro Moschitti	3005
<i>Syntax-BERT: Improving Pre-trained Transformers with Syntax Trees</i>	
Jiangang Bai, Yujing Wang, Yiren Chen, Yaming Yang, Jing Bai, Jing Yu and Yunhai Tong ..	3011
<i>DISK-CSV: Distilling Interpretable Semantic Knowledge with a Class Semantic Vector</i>	
Housam Khalifa Bashier, Mi-Young Kim and Randy Goebel	3021
<i>Attention Can Reflect Syntactic Structure (If You Let It)</i>	
Vinit Ravishankar, Artur Kulmizev, Mostafa Abdou, Anders Søgaard and Joakim Nivre	3031
<i>Extractive Summarization Considering Discourse and Coreference Relations based on Heterogeneous Graph</i>	
Yin Jou Huang and Sadao Kurohashi	3046
<i>CDA: a Cost Efficient Content-based Multilingual Web Document Aligner</i>	
Thuy Vu and Alessandro Moschitti	3053
<i>Metric-Type Identification for Multi-Level Header Numerical Tables in Scientific Papers</i>	
Lya Hulliyyatus Suadaa, Hidetaka Kamigaito, Manabu Okumura and Hiroya Takamura	3062
<i>EmpathBERT: A BERT-based Framework for Demographic-aware Empathy Prediction</i>	
Bhanu Prakash Reddy Guda, Aparna Garimella and Niyati Chhaya	3072
<i>Are Neural Networks Extracting Linguistic Properties or Memorizing Training Data? An Observation with a Multilingual Probe for Predicting Tense</i>	
Bingzhi Li and Guillaume Wisniewski	3080
<i>Is Supervised Syntactic Parsing Beneficial for Language Understanding Tasks? An Empirical Investigation</i>	
Goran Glavaš and Ivan Vulić	3090
<i>Facilitating Terminology Translation with Target Lemma Annotations</i>	
Toms Bergmanis and Mārcis Pinnis	3105
<i>Enhancing Sequence-to-Sequence Neural Lemmatization with External Resources</i>	
Kirill Milintsevich and Kairit Sirts	3112

<i>Summarising Historical Text in Modern Languages</i>	
Xutan Peng, Yi Zheng, Chenghua Lin and Advaith Siddharthan	3123
<i>Challenges in Automated Debiasing for Toxic Language Detection</i>	
Xuhui Zhou, Maarten Sap, Swabha Swayamdipta, Yejin Choi and Noah Smith	3143
<i>Adaptive Fusion Techniques for Multimodal Data</i>	
Gaurav Sahu and Olga Vechtomova	3156
<i>Detecting Scenes in Fiction: A new Segmentation Task</i>	
Albin Zehe, Leonard Konle, Lea Katharina Dümpelmann, Evelyn Gius, Andreas Hotho, Fotis Jannidis, Lucas Kaufmann, Markus Krug, Frank Puppe, Nils Reiter, Annekea Schreiber and Nathalie Wiedmer	3167
<i>LESA: Linguistic Encapsulation and Semantic Amalgamation Based Generalised Claim Detection from Online Content</i>	
Shreya Gupta, Parantak Singh, Megha Sundriyal, Md. Shad Akhtar and Tanmoy Chakraborty	3178
<i>Interpretability for Morphological Inflection: from Character-level Predictions to Subword-level Rules</i>	
Tatyana Ruzsics, Olga Sozinova, Ximena Gutierrez-Vasques and Tanja Samardzic	3189
<i>Expanding, Retrieving and Infilling: Diversifying Cross-Domain Question Generation with Flexible Templates</i>	
Xiaojing Yu and Anxiao Jiang	3202
<i>Handling Out-Of-Vocabulary Problem in Hangeul Word Embeddings</i>	
Ohjoon Kwon, Dohyun Kim, Soo-Ryeon Lee, Junyoung Choi and SangKeun Lee	3213
<i>Exploiting Multimodal Reinforcement Learning for Simultaneous Machine Translation</i>	
Julia Ive, Andy Mingren Li, Yishu Miao, Ozan Caglayan, Pranava Madhyastha and Lucia Specia	3222
<i>STAR: Cross-modal [STA]tement [R]epresentation for selecting relevant mathematical premises</i>	
Deborah Ferreira and André Freitas	3234
<i>Do Multi-Hop Question Answering Systems Know How to Answer the Single-Hop Sub-Questions?</i>	
Yixuan Tang, Hwee Tou Ng and Anthony Tung	3244
<i>Multilingual LAMA: Investigating Knowledge in Multilingual Pretrained Language Models</i>	
Nora Kassner, Philipp Dufter and Hinrich Schütze	3250
<i>Variational Weakly Supervised Sentiment Analysis with Posterior Regularization</i>	
Ziqian Zeng and Yangqiu Song	3259
<i>Framing Word Sense Disambiguation as a Multi-Label Problem for Model-Agnostic Knowledge Integration</i>	
Simone Conia and Roberto Navigli	3269
<i>Graph-based Fake News Detection using a Summarization Technique</i>	
Gihwan Kim and Youngjoong Ko	3276
<i>Cognition-aware Cognate Detection</i>	
Diptesh Kanojia, Prashant Sharma, Sayali Ghodekar, Pushpak Bhattacharyya, Gholamreza Haffari and malhar kulkarni	3281

<i>A Simple Three-Step Approach for the Automatic Detection of Exaggerated Statements in Health Science News</i>	
Jasabanta Patro and Sabyasachee Baruah	3293
<i>Modeling Coreference Relations in Visual Dialog</i>	
Mingxiao Li and Marie-Francine Moens	3306
<i>Increasing Robustness to Spurious Correlations using Forgettable Examples</i>	
Yadollah Yaghoobzadeh, Soroush Mehri, Remi Tachet des Combes, T. J. Hazen and Alessandro Sordoni	3319
<i>On Robustness of Neural Semantic Parsers</i>	
shuo huang, Zhuang Li, Lizhen Qu and Lei Pan	3333
<i>Benchmarking a transformer-FREE model for ad-hoc retrieval</i>	
Tiago Almeida and Sérgio Matos	3343
<i>Reanalyzing the Most Probable Sentence Problem: A Case Study in Explicating the Role of Entropy in Algorithmic Complexity</i>	
Eric Corlett and Gerald Penn	3354
<i>Probing the Probing Paradigm: Does Probing Accuracy Entail Task Relevance?</i>	
Abhilasha Ravichander, Yonatan Belinkov and Eduard Hovy	3363
<i>One-class Text Classification with Multi-modal Deep Support Vector Data Description</i>	
Chenlong Hu, Yukun Feng, Hidetaka Kamigaito, Hiroya Takamura and Manabu Okumura ...	3378
<i>Unsupervised Word Polysemy Quantification with Multiresolution Grids of Contextual Embeddings</i>	
Christos Xypolopoulos, Antoine Tixier and Michalis Vazirgiannis	3391
<i>Mega-COV: A Billion-Scale Dataset of 100+ Languages for COVID-19</i>	
Muhammad Abdul-Mageed, AbdelRahim Elmadany, El Moatez Billah Nagoudi, Dinesh Pabbi, Kunal Verma and Rannie Lin	3402
<i>Disfluency Correction using Unsupervised and Semi-supervised Learning</i>	
Nikhil Saini, Drumil Trivedi, Shreya Khare, Tejas Dhamecha, Preethi Jyothi, Samarth Bharadwaj and Pushpak Bhattacharyya	3421
<i>Complex Question Answering on knowledge graphs using machine translation and multi-task learning</i>	
Saurabh Srivastava, Mayur Patidar, Sudip Chowdhury, Puneet Agarwal, Indrajit Bhattacharya and Gautam Shroff	3428
<i>Recipes for Adapting Pre-trained Monolingual and Multilingual Models to Machine Translation</i>	
Asa Cooper Stickland, Xian Li and Marjan Ghazvininejad	3440
<i>From characters to words: the turning point of BPE merges</i>	
Ximena Gutierrez-Vasques, Christian Bentz, Olga Sozinova and Tanja Samardzic	3454
<i>A Large-scale Evaluation of Neural Machine Transliteration for Indic Languages</i>	
Anoop Kunchukuttan, Siddharth Jain and Rahul Kejriwal	3469
<i>Communicative-Function-Based Sentence Classification for Construction of an Academic Formulaic Expression Database</i>	
Kenichi Iwatsuki and Akiko Aizawa	3476

<i>Regulatory Compliance through Doc2Doc Information Retrieval: A case study in EU/UK legislation where text similarity has limitations</i>	
Ilias Chalkidis, Manos Fergadiotis, Nikolaos Manginas, Eva Katakalous and Prodromos Malakasiotis	3498
<i>The Chinese Remainder Theorem for Compact, Task-Precise, Efficient and Secure Word Embeddings</i>	
Patricia Thaine and Gerald Penn	3512
<i>Don't Change Me! User-Controllable Selective Paraphrase Generation</i>	
Mohan Zhang, Luchen Tan, Zihang Fu, Kun Xiong, Jimmy Lin, Ming Li and Zhengkai Tu	3522
<i>Rethinking Coherence Modeling: Synthetic vs. Downstream Tasks</i>	
Tasnim Mohiuddin, Prathyusha Jwalapuram, Xiang Lin and Shafiq Joty	3528
<i>From the Stage to the Audience: Propaganda on Reddit</i>	
Oana Balalau and Roxana Horincar	3540
<i>Probing for idiomaticity in vector space models</i>	
Marcos Garcia, Tiago Kramer Vieira, Carolina Scarton, Marco Idiart and Aline Villavicencio	3551
<i>Is the Understanding of Explicit Discourse Relations Required in Machine Reading Comprehension?</i>	
Yulong Wu, Viktor Schlegel and Riza Batista-Navarro	3565
<i>Why Is MBTI Personality Detection from Texts a Difficult Task?</i>	
Sanja Stajner and Seren Yenikent	3580
<i>Enconter: Entity Constrained Progressive Sequence Generation via Insertion-based Transformer</i>	
Lee Hsun Hsieh, Yang-Yin Lee and Ee-Peng Lim	3590
<i>Meta-Learning for Effective Multi-task and Multilingual Modelling</i>	
Ishan Tarunesh, Sushil Khyalia, vishwajeet kumar, Ganesh Ramakrishnan and Preethi Jyothi	3600
<i>"Killing Me" Is Not a Spoiler: Spoiler Detection Model using Graph Neural Networks with Dependency Relation-Aware Attention Mechanism</i>	
Buru Chang, Inggeol Lee, Hyunjae Kim and Jaewoo Kang	3613
<i>BERTese: Learning to Speak to BERT</i>	
Adi Haviv, Jonathan Berant and Amir Globerson	3618
<i>Lifelong Knowledge-Enriched Social Event Representation Learning</i>	
Prashanth Vijayaraghavan and Deb Roy	3624
<i>GLaRA: Graph-based Labeling Rule Augmentation for Weakly Supervised Named Entity Recognition</i>	
Xinyan Zhao, Haibo Ding and Zhe Feng	3636
<i>An End-to-end Model for Entity-level Relation Extraction using Multi-instance Learning</i>	
Markus Eberts and Adrian Ulges	3650
<i>WER-BERT: Automatic WER Estimation with BERT in a Balanced Ordinal Classification Paradigm</i>	
Akshay Krishna Sheshadri, Anvesh Rao Vijjini and Sukhdeep Kharbanda	3661
<i>Two Training Strategies for Improving Relation Extraction over Universal Graph</i>	
Qin Dai, Naoya Inoue, Ryo Takahashi and Kentaro Inui	3673
<i>Adaptation of Back-translation to Automatic Post-Editing for Synthetic Data Generation</i>	
WonKee Lee, Baikjin Jung, Jaehun Shin and Jong-Hyeok Lee	3685

<i>Removing Word-Level Spurious Alignment between Images and Pseudo-Captions in Unsupervised Image Captioning</i>	
Ukyo Honda, Yoshitaka Ushiku, Atsushi Hashimoto, Taro Watanabe and Yuji Matsumoto . . .	3692
<i>Towards More Fine-grained and Reliable NLP Performance Prediction</i>	
Zihuiwen Ye, Pengfei Liu, Jinlan Fu and Graham Neubig	3703
<i>Metrical Tagging in the Wild: Building and Annotating Poetry Corpora with Rhythmic Features</i>	
Thomas Haider	3715
<i>Enhancing Aspect-level Sentiment Analysis with Word Dependencies</i>	
Yuanhe Tian, Guimin Chen and Yan Song	3726