17th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2023)

Dubrovnik, Croatia
2-6 May 2023

Volume 1 of 5

ISBN: 978-1-7138-7375-4

Printed from e-media with permission by:

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

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

Copyright© (2023) by the Association for Computational Linguistics
All rights reserved.

Printed with permission by Curran Associates, Inc. (2023)

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

PiC: A Phrase-in-Context Dataset for Phrase Understanding and Semantic Search
Thang Pham, Seunghyun Yoon, Trung Bui and Anh Nguyen 1

Enhancing Dialogue Summarization with Topic-Aware Global- and Local- Level Centrality
Xinnian Liang, Shuangzhi Wu, Chenhao Cui, Jiaqi Bai, Chao Bian and Zhoujun Li 27

Exploiting Summarization Data to Help Text Simplification
Renliang Sun, Zhixian Yang and Xiaojun Wan 39

Shironaam: Bengali News Headline Generation using Auxiliary Information
Abu Ubaida Akash, Mir Tafseer Nayeem, Faisal Tareque Shohan and Tanvir Islam 52

PCC: Paraphrasing with Bottom-k Sampling and Cyclic Learning for Curriculum Data Augmentation
Hongyuan Lu and Wai Lam 68

A Two-Sided Discussion of Preregistration of NLP Research
Anders Søgaard, Daniel Hershcovich and Miryam de Lhoneux 83

WinoDict: Probing language models for in-context word acquisition
Julian Martin Eisenschlos, Jeremy R. Cole, Fangyu Liu and William W. Cohen 94

Sentiment as an Ordinal Latent Variable
Niklas Stoehr, Ryan Cotterell and Aaron Schein 103

Nationality Bias in Text Generation
Pranav Narayanan Venkit, Sanjana Gautam, Ruchi Panchanadikar, Ting-Hao Huang and Shomir Wilson 116

Investigating data partitioning strategies for crosslinguistic low-resource ASR evaluation
Zoey Liu, Justin Spence and Emily Prud'hommeaux 123

Shortcomings of Question Answering Based Factuality Frameworks for Error Localization
Ryo Kamoi, Tanya Goyal and Greg Durrett 132

Socratic Question Generation: A Novel Dataset, Models, and Evaluation
Beng Heng Ang, Sujatha Das Gollapalli and See-Kiong Ng 147

Do we need Label Regularization to Fine-tune Pre-trained Language Models?
Ivan Kobyzev, Aref Jafari, Mehdi Rezagholizadeh, Tianda Li, Alan Do-Omri, Peng Lu, Pascal Poupart and Ali Ghodsi 166

COVID-VTS: Fact Extraction and Verification on Short Video Platforms
Fuxiao Liu, Yaser Yacoob and Abhinav Shrivastava 178

Multimodal Graph Transformer for Multimodal Question Answering
Xuehai He and Xin Wang 189

Retrieval Enhanced Data Augmentation for Question Answering on Privacy Policies
Md Rizwan Parvez, Jianfeng Chi, Wasi Uddin Ahmad, Yuan Tian and Kai-Wei Chang 201

FastKASSIM: A Fast Tree Kernel-Based Syntactic Similarity Metric
Maximillian Chen, Caitlyn Chen, Xiao Yu and Zhou Yu 211

Friend-training: Learning from Models of Different but Related Tasks
Mian Zhang, Lifeng Jin, Linfeng Song, Haitao Mi, Xiabing Zhou and Dong Yu 232

Understanding Transformer Memorization Recall Through Idioms
Adi Haviv, Ido Cohen, Jacob Gidron, Roei Schuster, Yoav Goldberg and Mor Geva 248

A Discerning Several Thousand Judgments: GPT-3 Rates the Article + Adjective + Numeral + Noun Construction
Kyle Mahowald 265

Triple-Hybrid Energy-based Model Makes Better Calibrated Natural Language Understanding Models
Haotian Xu and Yingying Zhang 274

A weakly supervised textual entailment approach to zero-shot text classification
Marc Pàmies, Joan Llop, Francesco Multari, Nicolau Duran-Silva, César Parra-Rojas, Aitor Gonzalez-Agirre, Francesco Alessandro Massucci and Marta Villegas 286

Fair Enough: Standardizing Evaluation and Model Selection for Fairness Research in NLP
Xudong Han, Timothy Baldwin and Trevor Cohn 297

CHARD: Clinical Health-Aware Reasoning Across Dimensions for Text Generation Models
Steven Y. Feng, Vivek Khetan, Bogdan Sacaleanu, Anatole Gershman and Eduard Hovy 313

Prompt Tuning with Contradictory Intentions for Sarcasm Recognition
Yiyi Liu, Ruqing Zhang, Yixing Fan, Jiafeng Guo and Xueqi Cheng 328

COMBO: A Complete Benchmark for Open KG Canonicalization
Chengyue Jiang, Yong Jiang, Weiqi Wu, Yuting Zheng, Pengjun Xie and Kewei Tu 340

UScore: An Effective Approach to Fully Unsupervised Evaluation Metrics for Machine Translation
Jonas Belouadi and Steffen Eger 358

Assistive Recipe Editing through Critiquing
Diego Antognini, Shuyang Li, Boi Faltings and Julian McAuley 375

DiTTO: A Feature Representation Imitation Approach for Improving Cross-Lingual Transfer
Shanu Kumar, Soujanya Abbaraju, Sandipan Dandapat, Sunayana Sitaram and Monojit Choudhury 385

John is 50 years old, can his son be 65?Evaluating NLP Models' Understanding of Feasibility
Himanshu Gupta, Neeraj Varshney, Swaroop Mishra, Kuntal Kumar Pal, Saurabh Arjun Sawant, Kevin Scaria, Siddharth Goyal and Chitta Baral 407

Efficient Encoders for Streaming Sequence Tagging
Ayush Kaushal, Aditya Gupta, Shyam Upadhyay and Manaal Faruqui 418

Retrieve-and-Fill for Scenario-based Task-Oriented Semantic Parsing
Akshat Shrivastava, Shrey Desai, Anchit Gupta, Ali Elkahky, Aleksandr Livshits, Alexander Zotov and Ahmed Aly 430

Document Flattening: Beyond Concatenating Context for Document-Level Neural Machine Translation
Minghao Wu, George Foster, Lizhen Qu and Gholamreza Haffari 448

Scaling Back-Translation with Domain Text Generation for Sign Language Gloss Translation
Jinhui Ye, Wenxiang Jiao, Xing Wang and Zhaopeng Tu 463

Realistic Conversational Question Answering with Answer Selection based on Calibrated Confidence and Uncertainty Measurement
Soyeong Jeong, Jinheon Baek, Sung Ju Hwang and Jong Park 477

PANCETTA: Phoneme Aware Neural Completion to Elicit Tongue Twisters Automatically
Sedrick Scott Keh, Steven Y. Feng, Varun Gangal, Malihe Alikhani and Eduard Hovy 491

A User-Centered, Interactive, Human-in-the-Loop Topic Modelling System
Zheng Fang, Lama Alqazlan, Du Liu, Yulan He and Rob Procter 505

A Survey of Methods for Addressing Class Imbalance in Deep-Learning Based Natural Language Processing
Sophie Henning, William Beluch, Alexander Fraser and Annemarie Friedrich 523

Extracting or Guessing? Improving Faithfulness of Event Temporal Relation Extraction
Haoyu Wang, Hongming Zhang, Yuqian Deng, Jacob Gardner, Dan Roth and Muhao Chen . . 541

LoFT: Enhancing Faithfulness and Diversity for Table-to-Text Generation via Logic Form Control
Yilun Zhao, Zhenting Qi, Linyong Nan, Lorenzo Jaime Flores and Dragomir Radev 554

PromptDA: Label-guided Data Augmentation for Prompt-based Few Shot Learners
Canyu Chen and Kai Shu 562

Incorporating Question Answering-Based Signals into Abstractive Summarization via Salient Span Selection
Daniel Deutsch and Dan Roth 575

Patient Outcome and Zero-shot Diagnosis Prediction with Hypernetwork-guided Multitask Learning
Shaoxiong Ji and Pekka Marttinen 589

A Kind Introduction to Lexical and Grammatical Aspect, with a Survey of Computational Approaches
Annemarie Friedrich, Nianwen Xue and Alexis Palmer 599

Incorporating Context into Subword Vocabularies
Shaked Yehezkel and Yuval Pinter 623

LoRaLay: A Multilingual and Multimodal Dataset for Long Range and Layout-Aware Summarization
Laura Nguyen, Thomas Scialom, Benjamin Piwowarski and Jacopo Staiano 636

ViHOS: Hate Speech Spans Detection for Vietnamese
Phu Gia Hoang, Canh Duc Luu, Khanh Quoc Tran, Kiet Van Nguyen and Ngan Luu-Thuy Nguyen 652

Vote'n'Rank: Revision of Benchmarking with Social Choice Theory
Mark Rofin, Vladislav Mikhailov, Mikhail Florinsky, Andrey Kravchenko, Tatiana Shavrina, Elena Tutubalina, Daniel Karabekyan and Ekaterina Artemova 670

Combining Parameter-efficient Modules for Task-level Generalisation
Edoardo Maria Ponti, Alessandro Sordoni, Yoshua Bengio and Siva Reddy 687

Self-imitation Learning for Action Generation in Text-based Games
Zijing Shi, Yunqiu Xu, Meng Fang and Ling Chen 703

Investigating the Effect of Relative Positional Embeddings on AMR-to-Text Generation with Structural Adapters
Sebastien Montella, Alexis Nasr, Johannes Heinecke, Frederic Bechet and Lina M. Rojas Barahona 727

On the Intersection of Context-Free and Regular Languages
Clemente Pasti, Andreas Opedal, Tiago Pimentel, Tim Vieira, Jason Eisner and Ryan Cotterell 737

Social Influence Dialogue Systems: A Survey of Datasets and Models For Social Influence Tasks
Kushal Chawla, Weiyan Shi, Jingwen Zhang, Gale Lucas, Zhou Yu and Jonathan Gratch 750

Aggregating Crowdsourced and Automatic Judgments to Scale Up a Corpus of Anaphoric Reference for Fiction and Wikipedia Texts
Juntao Yu, Silviu Paun, Maris Camilleri, Paloma Garcia, Jon Chamberlain, Udo Kruschwitz and Massimo Poesio.....767

What Makes Sentences Semantically Related? A Textual Relatedness Dataset and Empirical Study
Mohamed Abdalla, Krishnapriya Vishnubhotla and Saif Mohammad.....782

RevUp: Revise and Update Information Bottleneck for Event Representation
Mehdi Rezaee and Francis Ferraro.....797

NusaX: Multilingual Parallel Sentiment Dataset for 10 Indonesian Local Languages
Genta Indra Winata, Alham Fikri Aji, Samuel Cahyawijaya, Rahmad Mahendra, Fajri Koto, Ade Romadhony, Kemal Kurniawan, David Moeljadi, Radityo Eko Prasojo, Pascale Fung, Timothy Baldwin, Jey Han Lau, Rico Sennrich and Sebastian Ruder.....815

The Functional Relevance of Probed Information: A Case Study
Michael Hanna, Roberto Zamparelli and David Mareček.....835

Do Pretrained Contextual Language Models Distinguish between Hebrew Homograph Analyses?
Avi Shmidman, Cheyn Shmuel Shmidman, Dan Bareket, Moshe Koppel and Reut Tsarfaty .. 849

Parameter-Efficient Tuning with Special Token Adaptation
Xiaocong Yang, James Y. Huang, Wenxuan Zhou and Muhao Chen.....865

Probing Power by Prompting: Harnessing Pre-trained Language Models for Power Connotation Framing
Shima Khanehzar, Trevor Cohn, Gosia Mikolajczak and Lea Frermann.....873

Zero and Few-Shot Localization of Task-Oriented Dialogue Agents with a Distilled Representation
Mehrad Moradshahi, Sina Semnani and Monica Lam.....886

Contextual Semantic Parsing for Multilingual Task-Oriented Dialogues
Mehrad Moradshahi, Victoria Tsai, Giovanni Campagna and Monica Lam.....902

Teacher Intervention: Improving Convergence of Quantization Aware Training for Ultra-Low Precision Transformers
Minsoo Kim, Kyuhong Shim, Seongmin Park, Wonyong Sung and Jungwook Choi.....916

Generative Replay Inspired by Hippocampal Memory Indexing for Continual Language Learning
Aru Maekawa, Hidetaka Kamigaito, Kotaro Funakoshi and Manabu Okumura.....930

A Survey of Multi-task Learning in Natural Language Processing: Regarding Task Relatedness and Training Methods
Zhihan Zhang, Wenhao Yu, Mengxia Yu, Zhichun Guo and Meng Jiang.....943

Conclusion-based Counter-Argument Generation
Milad Alshomary and Henning Wachsmuth.....957

Question-Answer Sentence Graph for Joint Modeling Answer Selection
Roshni Iyer, Thuy Vu, Alessandro Moschitti and Yizhou Sun.....968

Evaluating and Improving the Coreference Capabilities of Machine Translation Models
Asaf Yehudai, Arie Cattan, Omri Abend and Gabriel Stanovsky.....980

Document-Level Planning for Text Simplification
Liam Cripwell, Joël Legrand and Claire Gardent 993

Efficient Hybrid Generation Framework for Aspect-Based Sentiment Analysis
Haoran Lv, Junyi Liu, Henan Wang, Yaoming Wang, Jixiang Luo and Yaxiao Liu 1007

What's New? Summarizing Contributions in Scientific Literature
Hiroaki Hayashi, Wojciech Kryscinski, Bryan McCann, Nazneen Rajani and Caiming Xiong 1019

Find Parent then Label Children: A Two-stage Taxonomy Completion Method with Pre-trained Language Model
Fei Xia, Yixuan Weng, Shizhu He, Kang Liu and Jun Zhao 1032

Meta Self-Refinement for Robust Learning with Weak Supervision
Dawei Zhu, Xiaoyu Shen, Michael Hedderich and Dietrich Klakow 1043

Looking for a Needle in a Haystack: A Comprehensive Study of Hallucinations in Neural Machine Translation
Nuno M. Guerreiro, Elena Voita and André Martins 1059

Investigating UD Treebanks via Dataset Difficulty Measures
Artur Kulmizev and Joakim Nivre 1076

On Robustness of Prompt-based Semantic Parsing with Large Pre-trained Language Model: An Empirical Study on Codex
Terry Yue Zhuo, Zhuang Li, Yujin Huang, Fatemeh Shiri, Weiqing Wang, Gholamreza Haffari and Yuan-Fang Li 1090

Leveraging Task Dependency and Contrastive Learning for Case Outcome Classification on European Court of Human Rights Cases
Santosh T.Y.S.S, Marcel Perez San Blas, Phillip Kemper and Matthias Grabmair 1103

Semi-supervised Relation Extraction via Data Augmentation and Consistency-training
Komal Teru 1112

Event Temporal Relation Extraction with Bayesian Translational Model
Xingwei Tan, Gabriele Pergola and Yulan He 1125

Persona Expansion with Commonsense Knowledge for Diverse and Consistent Response Generation
Donghyun Kim, Youbin Ahn, Wongyu Kim, Chanhee Lee, Kyungchan Lee, Kyong-Ho Lee, Jeonguk Kim, Donghoon Shin and Yeonsoo Lee 1139

UnifEE: Unified Evidence Extraction for Fact Verification
Nan Hu, Zirui Wu, Yuxuan Lai, Chen Zhang and Yansong Feng 1150

MiniALBERT: Model Distillation via Parameter-Efficient Recursive Transformers
Mohammadmahdi Nouriborji, Omid Rohanian, Samaneh Kouchaki and David A. Clifton . . 1161

Multilingual Normalization of Temporal Expressions with Masked Language Models
Lukas Lange, Jannik Strötgen, Heike Adel and Dietrich Klakow 1174

K-hop neighbourhood regularization for few-shot learning on graphs: A case study of text classification
Niels van der Heijden, Ekaterina Shutova and Helen Yannakoudakis 1187

What Clued the AI Doctor In? On the Influence of Data Source and Quality for Transformer-Based Medical Self-Disclosure Detection
Mina Valizadeh, Xing Qian, Pardis Ranjbar-Noiey, Cornelia Caragea and Natalie Parde 1201

Improving Visual-Semantic Embedding with Adaptive Pooling and Optimization Objective
Zijian Zhang, Chang Shu, Ya Xiao, Yuan Shen, Di Zhu, Youxin Chen, Jing Xiao, Jey Han Lau, Qian Zhang and Zheng Lu 1217

Policy-based Reinforcement Learning for Generalisation in Interactive Text-based Environments
Edan Toledo, Jan Buys and Jonathan Shock 1230

Logic Against Bias: Textual Entailment Mitigates Stereotypical Sentence Reasoning
Hongyin Luo and James Glass 1243

Entity Tracking via Effective Use of Multi-Task Learning Model and Mention-guided Decoding
Janvijay Singh, Fan Bai and Zhen Wang 1255

Conversational Tree Search: A New Hybrid Dialog Task
Dirk Väth, Lindsey Vanderlyn and Ngoc Thang Vu 1264

A Human Subject Study of Named Entity Recognition in Conversational Music Recommendation Queries
Elena Epure and Romain Hennequin 1281

Entity Disambiguation with Entity Definitions
Luigi Procopio, Simone Conia, Edoardo Barba and Roberto Navigli 1297

Exploring Paracrawl for Document-level Neural Machine Translation
Yusser Al Ghussin, Jingyi Zhang and Josef van Genabith 1304

Poor Man's Quality Estimation: Predicting Reference-Based MT Metrics Without the Reference
Vilém Zouhar, Shehzaad Dhuliawala, Wangchunshu Zhou, Nico Daheim, Tom Kocmi, Yuchen Eleanor Jiang and Mrinmaya Sachan 1311

Integrating Translation Memories into Non-Autoregressive Machine Translation
Jitao Xu, Josep Crego and François Yvon 1326

Shorten the Long Tail for Rare Entity and Event Extraction
Pengfei Yu and Heng Ji 1339

Do Deep Neural Networks Capture Compositionality in Arithmetic Reasoning?
Keito Kudo, Yoichi Aoki, Tatsuki Kuribayashi, Ana Brassard, Masashi Yoshikawa, Keisuke Sakaguchi and Kentaro Inui 1351

BLM-AgrF: A New French Benchmark to Investigate Generalization of Agreement in Neural Networks
Aixiu An, Chunyang Jiang, Maria A. Rodriguez, Vivi Nastase and Paola Merlo 1363

Robustification of Multilingual Language Models to Real-world Noise in Crosslingual Zero-shot Settings with Robust Contrastive Pretraining
Asa Cooper Stickland, Sailik Sengupta, Jason Krone, Saab Mansour and He He 1375

Unsupervised Anomaly Detection in Multi-Topic Short-Text Corpora
Mira Ait-Saada and Mohamed Nadif 1392

Metaphor Detection with Effective Context Denoising
Shun Wang, Yucheng Li, Chenghua Lin, Loic Barrault and Frank Guerin 1404

Low-Resource Compositional Semantic Parsing with Concept Pretraining
Subendhu Rongali, Mukund Sridhar, Haidar Khan, Konstantine Arkoudas, Wael Hamza and Andrew McCallum 1410

Made of Steel? Learning Plausible Materials for Components in the Vehicle Repair Domain
Annerose Eichel, Helena Schlipf and Sabine Schulte im Walde 1420

Self-Adapted Utterance Selection for Suicidal Ideation Detection in Lifeline Conversations
Zhong-Ling Wang, Po-Hsien Huang, Wen-Yau Hsu and Hen-Hsen Huang 1436

Can Pretrained Language Models (Yet) Reason Deductively?
Zhangdie Yuan, Songbo Hu, Ivan Vulić, Anna Korhonen and Zaiqiao Meng 1447

Selective In-Context Data Augmentation for Intent Detection using Pointwise V-Information
Yen-Ting Lin, Alexandros Papangelis, Seokhwan Kim, Sungjin Lee, Devamanyu Hazarika, Mahdi Namazifar, Di Jin, Yang Liu and Dilek Hakkani-Tur 1463

Multilingual Representation Distillation with Contrastive Learning
Weiting Tan, Kevin Heffernan, Holger Schwenk and Philipp Koehn 1477

On the inconsistency of separable losses for structured prediction
Caio Corro 1491

A Systematic Search for Compound Semantics in Pretrained BERT Architectures
Filip Miletic and Sabine Schulte im Walde 1499

Efficiently Upgrading Multilingual Machine Translation Models to Support More Languages
Simeng Sun, Maha Elbayad, Anna Sun and James Cross 1513

Summarize and Generate to Back-translate: Unsupervised Translation of Programming Languages
Wasi Uddin Ahmad, Saikat Chakraborty, Baishakhi Ray and Kai-Wei Chang 1528

The Impacts of Unanswerable Questions on the Robustness of Machine Reading Comprehension Models
Son Quoc Tran, Phong Nguyen-Thuan Do, Uyen Le and Matt Kretchmar 1543

FrameBERT: Conceptual Metaphor Detection with Frame Embedding Learning
Yucheng Li, Shun Wang, Chenghua Lin, Frank Guerin and Loic Barrault 1558

Towards More Efficient Insertion Transformer with Fractional Positional Encoding
Zhisong Zhang, Yizhe Zhang and Bill Dolan 1564

SODAPOP: Open-Ended Discovery of Social Biases in Social Commonsense Reasoning Models
Haozhe An, Zongxia Li, Jieyu Zhao and Rachel Rudinger 1573

Augmenting Pre-trained Language Models with QA-Memory for Open-Domain Question Answering
Wenhu Chen, Pat Verga, Michiel de Jong, John Wieting and William W. Cohen 1597

Gold Doesn't Always Glitter: Spectral Removal of Linear and Nonlinear Guarded Attribute Information
Shun Shao, Yftah Ziser and Shay B. Cohen 1611

CTC Alignments Improve Autoregressive Translation
Brian Yan, Siddharth Dalmia, Yosuke Higuchi, Graham Neubig, Florian Metze, Alan W Black and Shinji Watanabe 1623

Modelling Temporal Document Sequences for Clinical ICD Coding
Boon Liang Clarence Ng, Diogo Santos and Marek Rei 1640

LongEval: Guidelines for Human Evaluation of Faithfulness in Long-form Summarization
Kalpesh Krishna, Erin Bransom, Bailey Kuehl, Mohit Iyyer, Pradeep Dasigi, Arman Cohan and Kyle Lo 1650

Cluster-Guided Label Generation in Extreme Multi-Label Classification
Taehee Jung, Joo-kyung Kim, Sungjin Lee and Dongyeop Kang 1670

Empathy Identification Systems are not Accurately Accounting for Context
Andrew Lee, Jonathan Kummerfeld, Larry An and Rada Mihalcea 1686

Enhancing Multi-Document Summarization with Cross-Document Graph-based Information Extraction
Zixuan Zhang, Heba Elfardy, Markus Dreyer, Kevin Small, Heng Ji and Mohit Bansal 1696

What happens before and after: Multi-Event Commonsense in Event Coreference Resolution
Sahithya Ravi, Chris Tanner, Raymond Ng and Vered Shwartz 1708

Multi-Modal Bias: Introducing a Framework for Stereotypical Bias Assessment beyond Gender and Race in Vision–Language Models
Sepehr Janghorbani and Gerard De Melo 1725

CylE: Cylinder Embeddings for Multi-hop Reasoning over Knowledge Graphs
Chau Duc Minh Nguyen, Tim French, Wei Liu and Michael Stewart 1736

Fiction-Writing Mode: An Effective Control for Human-Machine Collaborative Writing
Wenjie Zhong, Jason Naradowsky, Hiroya Takamura, Ichiro Kobayashi and Yusuke Miyao . 1752

Robustness Challenges in Model Distillation and Pruning for Natural Language Understanding
Mengnan Du, Subhabrata Mukherjee, Yu Cheng, Milad Shokouhi, Xia Hu and Ahmed Hassan Awadallah 1766

Don't Blame the Annotator: Bias Already Starts in the Annotation Instructions
Mihir Parmar, Swaroop Mishra, Mor Geva and Chitta Baral 1779

Performance Prediction via Bayesian Matrix Factorisation for Multilingual Natural Language Processing Tasks
Viktoria Schram, Daniel Beck and Trevor Cohn 1790

Unified Neural Topic Model via Contrastive Learning and Term Weighting
Sungwon Han, Mingi Shin, Sungkyu Park, Changwook Jung and Meeyoung Cha 1802

Don't Mess with Mister-in-Between: Improved Negative Search for Knowledge Graph Completion
Fan Jiang, Tom Drummond and Trevor Cohn 1818

Semantic Frame Induction with Deep Metric Learning
Kosuke Yamada, Ryohei Sasano and Koichi Takeda 1833

The Devil is in the Details: On Models and Training Regimes for Few-Shot Intent Classification
Mohsen Mesgar, Thy Thy Tran, Goran Glavaš and Iryna Gurevych 1846

Iterative Document-level Information Extraction via Imitation Learning
Yunmo Chen, William Gantt, Weiwei Gu, Tongfei Chen, Aaron White and Benjamin Van Durme 1858

CLICK: Contrastive Learning for Injecting Contextual Knowledge to Conversational Recommender System
Hyeongjun Yang, Heesoo Won, Youbin Ahn and Kyong-Ho Lee 1875

LEALLA: Learning Lightweight Language-agnostic Sentence Embeddings with Knowledge Distillation
Zhuoyuan Mao and Tetsuji Nakagawa 1886

Synthesizing Human Gaze Feedback for Improved NLP Performance
Varun Khurana, Yaman Kumar, Nora Hollenstein, Rajesh Kumar and Balaji Krishnamurthy 1895

Memory-efficient Temporal Moment Localization in Long Videos
Cristian Rodriguez-Opazo, Edison Marrese-Taylor, Basura Fernando, Hiroya Takamura and Qi Wu 1909

Extracting Victim Counts from Text
Mian Zhong, Shehzaad Dhuliawala and Niklas Stoehr 1925

ConEntail: An Entailment-based Framework for Universal Zero and Few Shot Classification with Supervised Contrastive Pretraining
Ranran Haoran Zhang, Aysa Xuemo Fan and Rui Zhang 1941

Guide the Learner: Controlling Product of Experts Debiasing Method Based on Token Attribution Similarities
Ali Modarressi, Hossein Amirkhani and Mohammad Taher Pilehvar 1954

Task and Sentiment Adaptation for Appraisal Tagging
Lin Tian, Xiuzhen Zhang, Myung Hee Kim and Jennifer Biggs 1960

DREEAM: Guiding Attention with Evidence for Improving Document-Level Relation Extraction
Youmi Ma, An Wang and Naoaki Okazaki 1971

Span-based Named Entity Recognition by Generating and Compressing Information
Nhung T. H. Nguyen, Makoto Miwa and Sophia Ananiadou 1984

An In-depth Analysis of Implicit and Subtle Hate Speech Messages
Nicolas Ocampo, Ekaterina Sviridova, Elena Cabrio and Serena Villata 1997

MTEB: Massive Text Embedding Benchmark
Niklas Muennighoff, Nouamane Tazi, Loic Magne and Nils Reimers 2014

Step by Step Loss Goes Very Far: Multi-Step Quantization for Adversarial Text Attacks
Piotr Gaiński and Klaudia Bałazy 2038

TwiRGCN: Temporally Weighted Graph Convolution for Question Answering over Temporal Knowledge Graphs
Aditya Sharma, Apoorv Saxena, Chitrank Gupta, Mehran Kazemi, Partha Talukdar and Soumen Chakrabarti 2049

ZELDA: A Comprehensive Benchmark for Supervised Entity Disambiguation
Marcel Milich and Alan Akbik 2061

GLADIS: A General and Large Acronym Disambiguation Benchmark
Lihu Chen, Gael Varoquaux and Fabian M. Suchanek 2073

Probing Cross-Lingual Lexical Knowledge from Multilingual Sentence Encoders
Ivan Vulić, Goran Glavaš, Fangyu Liu, Nigel Collier, Edoardo Maria Ponti and Anna Korhonen 2089

Pento-DIARef: A Diagnostic Dataset for Learning the Incremental Algorithm for Referring Expression Generation from Examples
Philipp Sadler and David Schlangen 2106

Mitigating Exposure Bias in Grammatical Error Correction with Data Augmentation and Reweighting
Hannan Cao, Wenmian Yang and Hwee Tou Ng 2123

Plausible May Not Be Faithful: Probing Object Hallucination in Vision-Language Pre-training
Wenliang Dai, Zihan Liu, Ziwei Ji, Dan Su and Pascale Fung 2136

Characterizing the Entities in Harmful Memes: Who is the Hero, the Villain, the Victim?
Shivam Sharma, Atharva Kulkarni, Tharun Suresh, Himanshi Mathur, Preslav Nakov, Md. Shad Akhtar and Tanmoy Chakraborty 2149

Systematic Investigation of Strategies Tailored for Low-Resource Settings for Low-Resource Dependency Parsing
Jivnesh Sandhan, Laxmidhar Behera and Pawan Goyal 2164

Compositional Generalisation with Structured Reordering and Fertility Layers
Matthias Lindemann, Alexander Koller and Ivan Titov 2172

Investigating Multi-source Active Learning for Natural Language Inference
Ard Snijders, Douwe Kiela and Katerina Margatina 2187

Towards a Unified Multi-Domain Multilingual Named Entity Recognition Model
Mayank Kulkarni, Daniel Preotiuc-Pietro, Karthik Radhakrishnan, Genta Indra Winata, Shijie Wu, Lingjue Xie and Shaohua Yang 2210

Do Neural Topic Models Really Need Dropout? Analysis of the Effect of Dropout in Topic Modeling
Suman Adhya, Avishek Lahiri and Debarshi Kumar Sanyal 2220

A Psycholinguistic Analysis of BERT's Representations of Compounds
Lars Buijtelaar and Sandro Pezzelle 2230

Measuring Normative and Descriptive Biases in Language Models Using Census Data
Samia Touileb, Lilja Øvrelid and Erik Velldal 2242

UDAPTER - Efficient Domain Adaptation Using Adapters
Bhavitvya Malik, Abhinav Ramesh Kashyap, Min-Yen Kan and Soujanya Poria 2249

Efficient CTC Regularization via Coarse Labels for End-to-End Speech Translation
Biao Zhang, Barry Haddow and Rico Sennrich 2264

Exploring Category Structure with Contextual Language Models and Lexical Semantic Networks
Joseph Renner, Pascal Denis, Remi Gilleron and Angèle Brunellière 2277

An Empirical Study of Clinical Note Generation from Doctor-Patient Encounters
Asma Ben Abacha, Wen-wai Yim, Yadan Fan and Thomas Lin 2291

Instruction Clarification Requests in Multimodal Collaborative Dialogue Games: Tasks, and an Analysis of the CoDraw Dataset
Brielen Madureira and David Schlangen 2303

Can Synthetic Text Help Clinical Named Entity Recognition? A Study of Electronic Health Records in French
Nicolas Hiebel, Olivier Ferret, Karen Fort and Aurélie Névéol 2320

IRMA: the 335-million-word Italian coRpus for studying MisinformAtion
Fabio Carrella, Alessandro Miani and Stephan Lewandowsky 2339

Parameter-Efficient Korean Character-Level Language Modeling
Marco Cognetta, Sangwhan Moon, Lawrence Wolf-sonkin and Naoaki Okazaki 2350

Opportunities and Challenges in Neural Dialog Tutoring
Jakub Macina, Nico Daheim, Lingzhi Wang, Tanmay Sinha, Manu Kapur, Iryna Gurevych and Mrinmaya Sachan 2357

Evaluating the Robustness of Discrete Prompts
Yoichi Ishibashi, Danushka Bollegala, Katsuhito Sudoh and Satoshi Nakamura 2373

Assessing Out-of-Domain Language Model Performance from Few Examples
Prasann Singhal, Jarad Forristal, Xi Ye and Greg Durrett 2385

Mind the Labels: Describing Relations in Knowledge Graphs With Pretrained Models
Zdeněk Kasner, Ioannis Konstas and Ondrej Dusek 2398

Shapley Head Pruning: Identifying and Removing Interference in Multilingual Transformers
William Held and Diyi Yang 2416

Why Don't You Do It Right? Analysing Annotators' Disagreement in Subjective Tasks
Marta Sandri, Elisa Leonardelli, Sara Tonelli and Elisabetta Jezek 2428

Analyzing Challenges in Neural Machine Translation for Software Localization
Sai Koneru, Matthias Huck, Miriam Exel and Jan Niehues 2442

Bootstrapping Multilingual Semantic Parsers using Large Language Models
Abhijeet Awasthi, Nitish Gupta, Bidisha Samanta, Shachi Dave, Sunita Sarawagi and Partha Talukdar 2455

Modeling Complex Event Scenarios via Simple Entity-focused Questions
Mahnaz Koupaee, Greg Durrett, Nathanael Chambers and Niranjan Balasubramanian 2468

Uncovering Implicit Inferences for Improved Relational Argument Mining
Ameer Saadat-Yazdi, Jeff Z. Pan and Nadin Kokciyan 2484

How people talk about each other: Modeling Generalized Intergroup Bias and Emotion
Venkata Subrahmanyan Govindarajan, Katherine Atwell, Barea Sinno, Malihe Alikhani, David I. Beaver and Junyi Jessy Li 2496

Semantic Parsing for Conversational Question Answering over Knowledge Graphs
Laura Perez-Beltrachini, Parag Jain, Emilio Monti and Mirella Lapata 2507

MAPL: Parameter-Efficient Adaptation of Unimodal Pre-Trained Models for Vision-Language Few-Shot Prompting
Oscar Mañas, Pau Rodriguez Lopez, Saba Ahmadi, Aida Nematzadeh, Yash Goyal and Aishwarya Agrawal 2523

ComSearch: Equation Searching with Combinatorial Strategy for Solving Math Word Problems with Weak Supervision
Qianying Liu, Wenyu Guan, Jianhao Shen, Fei Cheng and Sadao Kurohashi 2549

Towards preserving word order importance through Forced Invalidation
Hadeel Al-Negheimish, Pranava Madhyastha and Alessandra Russo 2563

How Many and Which Training Points Would Need to be Removed to Flip this Prediction?
Jinghan Yang, Sarthak Jain and Byron C. Wallace 2571

Reinforced Sequence Training based Subjective Bias Correction
Karthic Madanagopal and James Caverlee 2585

Detecting Lexical Borrowings from Dominant Languages in Multilingual Wordlists
John E. Miller and Johann-Mattis List 2599

Towards Integration of Discriminability and Robustness for Document-Level Relation Extraction
Jia Guo, Stanley Kok and Lidong Bing 2606

Penguins Don't Fly: Reasoning about Generics through Instantiations and Exceptions
Emily Allaway, Jena D. Hwang, Chandra Bhagavatula, Kathleen McKeown, Doug Downey and Yejin Choi 2618

Adding Instructions during Pretraining: Effective way of Controlling Toxicity in Language Models
Shrimai Prabhumoye, Mostofa Patwary, Mohammad Shoeybi and Bryan Catanzaro 2636

Multi2Claim: Generating Scientific Claims from Multi-Choice Questions for Scientific Fact-Checking
Neset Tan, Trung Nguyen, Josh Bensemann, Alex Peng, Qiming Bao, Yang Chen, Mark Gahegan and Michael Witbrock 2652

On Evaluation of Document Classifiers using RVL-CDIP
Stefan Larson, Gordon Lim and Kevin Leach 2665

Event Linking: Grounding Event Mentions to Wikipedia
Xiaodong Yu, Wenpeng Yin, Nitish Gupta and Dan Roth 2679

SwitchPrompt: Learning Domain-Specific Gated Soft Prompts for Classification in Low-Resource Domains
Koustava Goswami, Lukas Lange, Jun Araki and Heike Adel 2689

Do dialogue representations align with perception? An empirical study
Sarenne Wallbridge, Peter Bell and Catherine Lai 2696

Methods for Measuring, Updating, and Visualizing Factual Beliefs in Language Models
Peter Hase, Mona Diab, Asli Celikyilmaz, Xian Li, Zornitsa Kozareva, Veselin Stoyanov, Mohit Bansal and Srinivasan Iyer 2714

Improving Sign Recognition with Phonology
Lee Kezar, Jesse Thomason and Zed Sehyr 2732

Parameter-efficient Modularised Bias Mitigation via AdapterFusion
Deepak Kumar, Oleg Lesota, George Zerveas, Daniel Cohen, Carsten Eickhoff, Markus Schedl and Navid Rekabsaz 2738

LingMess: Linguistically Informed Multi Expert Scorers for Coreference Resolution
Shon Otmazgin, Arie Cattan and Yoav Goldberg 2752

Finding the Law: Enhancing Statutory Article Retrieval via Graph Neural Networks
Antoine Louis, Gijs van Dijck and Gerasimos Spanakis 2761

Behavior Cloned Transformers are Neurosymbolic Reasoners
Ruoyao Wang, Peter Jansen, Marc-Alexandre Côté and Prithviraj Ammanabrolu 2777

Bridging the Gap Between BabelNet and HowNet: Unsupervised Sense Alignment and Sememe Prediction
Xiang Zhang, Ning Shi, Bradley Hauer and Grzegorz Kondrak 2789

The StatCan Dialogue Dataset: Retrieving Data Tables through Conversations with Genuine Intents
Xing Han Lu, Siva Reddy and Harm de Vries 2799

Question Generation Using Sequence-to-Sequence Model with Semantic Role Labels
Alireza Naeiji, Aijun An, Heidar Davoudi, Marjan Delpisheh and Muath Alzghool 2830

StyLEx: Explaining Style Using Human Lexical Annotations
Shirley Anugrah Hayati, Kyumin Park, Dheeraj Rajagopal, Lyle Ungar and Dongyeop Kang 2843

Comparing Intrinsic Gender Bias Evaluation Measures without using Human Annotated Examples
Masahiro Kaneko, Danushka Bollegala and Naoaki Okazaki 2857

Faithfulness-Aware Decoding Strategies for Abstractive Summarization
David Wan, Mengwen Liu, Kathleen McKeown, Markus Dreyer and Mohit Bansal 2864

Dynamic Benchmarking of Masked Language Models on Temporal Concept Drift with Multiple Views
Katerina Margatina, Shuai Wang, Yogarshi Vyas, Neha Anna John, Yassine Benajiba and Miguel Ballesteros 2881

Real-Time Visual Feedback to Guide Benchmark Creation: A Human-and-Metric-in-the-Loop Workflow
Anjana Arunkumar, Swaroop Mishra, Bhavdeep Singh Sachdeva, Chitta Baral and Chris Bryan 2899

COMPS: Conceptual Minimal Pair Sentences for testing Robust Property Knowledge and its Inheritance in Pre-trained Language Models
Kanishka Misra, Julia Rayz and Allyson Ettinger 2928

Probabilistic Robustness for Data Filtering
Yu Yu, Abdul Rafae Khan, Shahram Khadivi and Jia Xu 2950

Unsupervised Improvement of Factual Knowledge in Language Models
Nafis Sadeq, Byungkyu Kang, Prarit Lamba and Julian McAuley 2960

Learning to Ignore Adversarial Attacks
Yiming Zhang, Yangqiaoyu Zhou, Samuel Carton and Chenhao Tan 2970

Should You Mask 15% in Masked Language Modeling?
Alexander Wettig, Tianyu Gao, Zexuan Zhong and Danqi Chen 2985

How do Words Contribute to Sentence Semantics? Revisiting Sentence Embeddings with a Perturbation Method
Wenlin Yao, Lifeng Jin, Hongming Zhang, Xiaoman Pan, Kaiqiang Song, Dian Yu, Dong Yu and Jianshu Chen 3001

AutoTriggER: Label-Efficient and Robust Named Entity Recognition with Auxiliary Trigger Extraction
Dong-Ho Lee, Ravi Kiran Selvam, Sheikh Muhammad Sarwar, Bill Yuchen Lin, Fred Morstatter, Jay Pujara, Elizabeth Boschee, James Allan and Xiang Ren 3011

Incorporating Task-Specific Concept Knowledge into Script Learning
Chenkai Sun, Tie Xu, ChengXiang Zhai and Heng Ji 3026

DeepMaven: Deep Question Answering on Long-Distance Movie/TV Show Videos with Multimedia Knowledge Extraction and Synthesis
Yi Fung, Han Wang, Tong Wang, Ali Kebarighotbi, Mohit Bansal, Heng Ji and Prem Natarajan 3041

Salient Span Masking for Temporal Understanding
Jeremy R. Cole, Aditi Chaudhary, Bhuwan Dhingra and Partha Talukdar 3052

PECO: Examining Single Sentence Label Leakage in Natural Language Inference Datasets through Progressive Evaluation of Cluster Outliers
Michael Saxon, Xinyi Wang, Wenda Xu and William Yang Wang 3061

Weakly-Supervised Questions for Zero-Shot Relation Extraction
Saeed Najafi and Alona Fyshe 3075

DiffQG: Generating Questions to Summarize Factual Changes
Jeremy R. Cole, Palak Jain, Julian Martin Eisenschlos, Michael J.Q. Zhang, Eunsol Choi and Bhuwan Dhingra 3088

Contextual Dynamic Prompting for Response Generation in Task-oriented Dialog Systems
Sandesh Swamy, Narges Tabari, Chacha Chen and Rashmi Gangadharaiah 3102

Why Can't Discourse Parsing Generalize? A Thorough Investigation of the Impact of Data Diversity
Yang Janet Liu and Amir Zeldes 3112

Enriching Biomedical Knowledge for Low-resource Language Through Large-scale Translation
Long Phan, Tai Dang, Hieu Tran, Trieu H. Trinh, Vy Phan, Lam D. Chau and Minh-Thang Luong 3131

Syntax-guided Neural Module Distillation to Probe Compositionality in Sentence Embeddings
Rohan Pandey 3143

Closed-book Question Generation via Contrastive Learning
Xiangjue Dong, Jiaying Lu, Jianling Wang and James Caverlee 3150

A Hybrid Detection and Generation Framework with Separate Encoders for Event Extraction
Ge Shi, Yunyue Su, Yongliang Ma and Ming Zhou 3163

Path Spuriousness-aware Reinforcement Learning for Multi-Hop Knowledge Graph Reasoning
Chunyang Jiang, Tianchen Zhu, Haoyi Zhou, Chang Liu, Ting Deng, Chunming Hu and Jianxin Li 3181

Self-Adaptive Named Entity Recognition by Retrieving Unstructured Knowledge
Kosuke Nishida, Naoki Yoshinaga and Kyosuke Nishida 3193

When Do Pre-Training Biases Propagate to Downstream Tasks? A Case Study in Text Summarization
Faisal Ladhak, Esin Durmus, Mirac Suzgun, Tianyi Zhang, Dan Jurafsky, Kathleen McKeown and Tatsunori Hashimoto 3206

BERT Shows Garden Path Effects
Tovah Irwin, Kyra Wilson and Alec Marantz 3220

Models Teaching Models: Improving Model Accuracy with Slingshot Learning
Lachlan O'Neill, Nandini Anantharama, Satya Borgohain and Simon D. Angus 3233

A Federated Approach for Hate Speech Detection
Jay Gala, Deep Gandhi, Jash Mehta and Zeerak Talat 3248

Learning the Legibility of Visual Text Perturbations
Dev Seth, Rickard Stureborg, Danish Pruthi and Bhuwan Dhingra 3260

DyLoRA: Parameter-Efficient Tuning of Pre-trained Models using Dynamic Search-Free Low-Rank Adaptation
Mojtaba Valipour, Mehdi Rezagholizadeh, Ivan Kobyzev and Ali Ghodsi 3274

Conversational Emotion-Cause Pair Extraction with Guided Mixture of Experts
DongJin Jeong and JinYeong Bak 3288

Language Generation Models Can Cause Harm: So What Can We Do About It? An Actionable Survey
Sachin Kumar, Vidhisha Balachandran, Lucille Njoo, Antonios Anastasopoulos and Yulia Tsvetkov 3299

TraVLR: Now You See It, Now You Don't! A Bimodal Dataset for Evaluating Visio-Linguistic Reasoning
Keng Ji Chow, Samson Tan and Min-Yen Kan 3322

Paraphrase Acquisition from Image Captions
Marcel Gohsen, Matthias Hagen, Martin Potthast and Benno Stein 3348

Generation-Based Data Augmentation for Offensive Language Detection: Is It Worth It?
Camilla Casula and Sara Tonelli 3359

Quantifying Context Mixing in Transformers
Hosein Mohebbi, Willem Zuidema, Grzegorz Chrupała and Afra Alishahi 3378

KGVL-BART: Knowledge Graph Augmented Visual Language BART for Radiology Report Generation
Kaveri Kale, Pushpak Bhattacharyya, Milind Gune, Aditya Shetty and Rustom Lawyer 3401

A simple but effective model for attachment in discourse parsing with multi-task learning for relation labeling
Zineb Bennis, Julie Hunter and Nicholas Asher 3412

How Far Can It Go? On Intrinsic Gender Bias Mitigation for Text Classification
Ewoenam Kwaku Tokpo, Pieter Delobelle, Bettina Berendt and Toon Calders 3418

Multimodal Event Transformer for Image-guided Story Ending Generation
Yucheng Zhou and Guodong Long 3434

Improving Cross-modal Alignment for Text-Guided Image Inpainting
Yucheng Zhou and Guodong Long 3445

Semantic Specialization for Knowledge-based Word Sense Disambiguation
Sakae Mizuki and Naoaki Okazaki 3457

Concept-based Persona Expansion for Improving Diversity of Persona-Grounded Dialogue
Donghyun Kim, Youbin Ahn, Chanhee Lee, Wongyu Kim, Kyong-Ho Lee, Donghoon Shin and Yeonsoo Lee 3471

RPTCS: A Reinforced Persona-aware Topic-guiding Conversational System
Zishan Ahmad, Kshitij Mishra, Asif Ekbal and Pushpak Bhattacharyya 3482

What Did You Learn To Hate? A Topic-Oriented Analysis of Generalization in Hate Speech Detection
Tom Bourgeade, Patricia Chiril, Farah Benamara and Véronique Moriceau 3495

End-to-end Case-Based Reasoning for Commonsense Knowledge Base Completion
Zonglin Yang, Xinya Du, Erik Cambria and Claire Cardie 3509

Exploring Segmentation Approaches for Neural Machine Translation of Code-Switched Egyptian Arabic-English Text
Marwa Gaser, Manuel Mager, Injy Hamed, Nizar Habash, Slim Abdennadher and Ngoc Thang Vu 3523

Identifying the limits of transformers when performing model-checking with natural language
Tharindu Madusanka, Riza Batista-navarro and Ian Pratt-hartmann 3539

Improving the Generalizability of Collaborative Dialogue Analysis With Multi-Feature Embeddings
Ayesha Enayet and Gita Sukthankar 3551

MetaQA: Combining Expert Agents for Multi-Skill Question Answering
Haritz Puerto, Gözde Şahin and Iryna Gurevych 3566

BERT Is Not The Count: Learning to Match Mathematical Statements with Proofs
Weixian Waylon Li, Yftah Ziser, Maximin Coavoux and Shay B. Cohen 3581

Lessons Learned from a Citizen Science Project for Natural Language Processing
Jan-Christoph Klie, Ji-Ung Lee, Kevin Stowe, Gözde Şahin, Nafise Sadat Moosavi, Luke Bates, Dominic Petrak, Richard Eckart De Castilho and Iryna Gurevych 3594

Contrastive Learning with Keyword-based Data Augmentation for Code Search and Code Question Answering
Shinwoo Park, Youngwook Kim and Yo-Sub Han 3609

Large Scale Multi-Lingual Multi-Modal Summarization Dataset
Yash Verma, Anubhav Jangra, Raghvendra Verma and Sriparna Saha 3620

External Knowledge Acquisition for End-to-End Document-Oriented Dialog Systems
Tuan M. Lai, Giuseppe Castellucci, Saar Kuzi, Heng Ji and Oleg Rokhlenko 3633

In-Depth Look at Word Filling Societal Bias Measures
Matúš Pikuliak, Ivana Beňová and Viktor Bachratý 3648

Retrieval-augmented Image Captioning
Rita Ramos, Desmond Elliott and Bruno Martins 3666

Automatic Evaluation and Analysis of Idioms in Neural Machine Translation
Christos Baziotis, Prashant Mathur and Eva Hasler 3682

Representation biases in sentence transformers
Dmitry Nikolaev and Sebastian Padó 3701

AbLit: A Resource for Analyzing and Generating Abridged Versions of English Literature
Melissa Roemmele, Kyle Shaffer, Katrina Olsen, Yiyi Wang and Steve DeNeefe 3717

Self-training Reduces Flicker in Retranslation-based Simultaneous Translation
Sukanta Sen, Rico Sennrich, Biao Zhang and Barry Haddow 3734

Social Commonsense for Explanation and Cultural Bias Discovery
Lisa Bauer, Hanna Tischer and Mohit Bansal 3745

Counter-GAP: Counterfactual Bias Evaluation through Gendered Ambiguous Pronouns
Zhongbin Xie, Vid Kocijan, Thomas Lukasiewicz and Oana-Maria Camburu 3761

The NLP Task Effectiveness of Long-Range Transformers
Guanghui Qin, Yukun Feng and Benjamin Van Durme 3774

Creation and evaluation of timelines for longitudinal user posts
Anthony Hills, Adam Tsakalidis, Federico Nanni, Ioannis Zachos and Maria Liakata 3791

Semi-supervised New Event Type Induction and Description via Contrastive Loss-Enforced Batch Attention
Carl Edwards and Heng Ji 3805

Multilingual Content Moderation: A Case Study on Reddit
Meng Ye, Karan Sikka, Katherine Atwell, Sabit Hassan, Ajay Divakaran and Malihe Alikhani 3828

GrIPS: Gradient-free, Edit-based Instruction Search for Prompting Large Language Models
Archiki Prasad, Peter Hase, Xiang Zhou and Mohit Bansal 3845

DiscoScore: Evaluating Text Generation with BERT and Discourse Coherence
Wei Zhao, Michael Strube and Steffen Eger 3865

Know your audience: specializing grounded language models with listener subtraction
Aaditya K Singh, David Ding, Andrew Saxe, Felix Hill and Andrew Lampinen 3884

Meeting the Needs of Low-Resource Languages: The Value of Automatic Alignments via Pretrained Models
Abteen Ebrahimi, Arya D. Mccarthy, Arturo Oncevay, John E. Ortega, Luis Chiruzzo, Gustavo Giménez-Lugo, Rolando Coto-Solano and Katharina Kann 3912