

Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)

[Chengqing Zong](#) | [Fei Xia](#) | [Wenjie Li](#) | [Roberto Navigli](#)

[pdfbibabs](#)

Investigating label suggestions for opinion mining in German Covid-19 social media

[Tilman Beck](#) | [Ji-Ung Lee](#) | [Christina Viehmann](#) | [Marcus Maurer](#) | [Oliver Quiring](#) | [Iryna Gurevych](#)

[pdfbibabs](#)

How Did This Get Funded?! Automatically Identifying Quirky Scientific Achievements

[Chen Shani](#) | [Nadav Borenstein](#) | [Dafna Shahaf](#)

[pdfbibabs](#)

Engage the Public: Poll Question Generation for Social Media Posts

[Zexin Lu](#) | [Keyang Ding](#) | [Yuji Zhang](#) | [Jing Li](#) | [Baolin Peng](#) | [Lemao Liu](#)

[pdfbibabs](#)

HateCheck: Functional Tests for Hate Speech Detection Models

[Paul Röttger](#) | [Bertie Vidgen](#) | [Dong Nguyen](#) | [Zeera Waseem](#) | [Helen Margetts](#) | [Janet Pierrehumbert](#)

[pdfbibabs](#)

Unified Dual-view Cognitive Model for Interpretable Claim Verification

[Lianwei Wu](#) | [Yuan Rao](#) | [Yuguan Lan](#) | [Ling Sun](#) | [Zhaoyin Qi](#)

[pdfbibabs](#)

DeepRapper: Neural Rap Generation with Rhyme and Rhythm Modeling

[Langqing Xue](#) | [Kaitao Song](#) | [Duocai Wu](#) | [Xu Tan](#) | [Nevin L. Zhang](#) | [Tao Qin](#) | [Wei-Qiang Zhang](#) | [Tie-Yan Liu](#)

[pdfbibabs](#)

PENS: A Dataset and Generic Framework for Personalized News Headline Generation

[Xiang Ao](#) | [Xiting Wang](#) | [Ling Luo](#) | [Ying Qiao](#) | [Qing He](#) | [Xing Xie](#)

[pdfbibabs](#)

Enhancing Content Preservation in Text Style Transfer Using Reverse Attention and Conditional Layer Normalization

[Dongkyu Lee](#) | [Zhiliang Tian](#) | [Langqing Xue](#) | [Nevin L. Zhang](#)

[pdfbibabs](#)

Mention Flags (MF): Constraining Transformer-based Text Generators

[Yufei Wang](#) | [Ian Wood](#) | [Stephen Wan](#) | [Mark Dras](#) | [Mark Johnson](#)

[pdfbibabs](#)

Generalising Multilingual Concept-to-Text NLG with Language Agnostic Delexicalisation

[Giulio Zhou](#) | [Gerasimos Lampouras](#)

[pdfbibabs](#)

Conversations Are Not Flat: Modeling the Dynamic Information Flow across Dialogue Utterances

[Zekang Li](#) | [Jinchao Zhang](#) | [Zhengcong Fei](#) | [Yang Feng](#) | [Jie Zhou](#)

[pdfbibabs](#)

Dual Slot Selector via Local Reliability Verification for Dialogue State Tracking

[Jinyu Guo](#) | [Kai Shuang](#) | [Jijie Li](#) | [Zihan Wang](#)

[pdfbibabs](#)

Transferable Dialogue Systems and User Simulators

[Bo-Hsiang Tseng](#) | [Yinpei Dai](#) | [Florian Kreyssig](#) | [Bill Byrne](#)

[pdfbibabs](#)

BoB: BERT Over BERT for Training Persona-based Dialogue Models from Limited Personalized Data

[Haoyu Song](#) | [Yan Wang](#) | [Kaiyan Zhang](#) | [Wei-Nan Zhang](#) | [Ting Liu](#)

[pdfbibabs](#)

GL-GIN: Fast and Accurate Non-Autoregressive Model for Joint Multiple Intent Detection and Slot Filling

[Libo Qin](#) | [Fuxuan Wei](#) | [Tianbao Xie](#) | [Xiao Xu](#) | [Wanxiang Che](#) | [Ting Liu](#)

[pdfbibabs](#)

Accelerating BERT Inference for Sequence Labeling via Early-Exit

[Xiaonan Li](#) | [Yunfan Shao](#) | [Tianxiang Sun](#) | [Hang Yan](#) | [Xipeng Qiu](#) | [Xuanjing Huang](#)

[pdfbibabs](#)

Modularized Interaction Network for Named Entity Recognition

[Fei Li](#) | [Zheng Wang](#) | [Siu Cheung Hui](#) | [Lejian Liao](#) | [Dandan Song](#) | [Jing Xu](#) | [Guoxiu He](#) | [Meihuizi Jia](#)

[pdfbibabs](#)

Capturing Event Argument Interaction via A Bi-Directional Entity-Level Recurrent Decoder

[Xi Xiangyu](#) | [Wei Ye](#) | [Shikun Zhang](#) | [Quanxiu Wang](#) | [Huixing Jiang](#) | [Wei Wu](#)

[pdfbibabs](#)

UniRE: A Unified Label Space for Entity Relation Extraction

[Yijun Wang](#) | [Changzhi Sun](#) | [Yuanbin Wu](#) | [Hao Zhou](#) | [Lei Li](#) | [Junchi Yan](#)

[pdfbibabs](#)

Refining Sample Embeddings with Relation Prototypes to Enhance Continual Relation Extraction

[Li Cui](#) | [Deqing Yang](#) | [Jiaxin Yu](#) | [Chengwei Hu](#) | [Jiayang Cheng](#) | [Jingjie Yi](#) | [Yanghua Xiao](#)

[pdfbibabs](#)

Contrastive Learning for Many-to-many Multilingual Neural Machine Translation

[Xiao Pan](#) | [Mingxuan Wang](#) | [Liwei Wu](#) | [Lei Li](#)

[pdfbibabs](#)

Understanding the Properties of Minimum Bayes Risk Decoding in Neural Machine Translation

[Mathias Müller](#) | [Rico Sennrich](#)

[pdfbibabs](#)

Multi-Head Highly Parallelized LSTM Decoder for Neural Machine Translation

[Hongfei Xu](#) | [Qiuhui Liu](#) | [Josef van Genabith](#) | [Deyi Xiong](#) | [Meng Zhang](#)

[pdfbibabs](#)

A Bidirectional Transformer Based Alignment Model for Unsupervised Word Alignment

[Jingyi Zhang](#) | [Josef van Genabith](#)

[pdfbibabs](#)

[Learning Language Specific Sub-network for Multilingual Machine Translation](#)

[Zehui Lin](#) | [Liwei Wu](#) | [Mingxuan Wang](#) | [Lei Li](#)

[pdfbibabs](#)

[Exploring the Efficacy of Automatically Generated Counterfactuals for Sentiment Analysis](#)

[Linyi Yang](#) | [Jiazheng Li](#) | [Padraig Cunningham](#) | [Yue Zhang](#) | [Barry Smyth](#) | [Ruihai Dong](#)

[pdfbibabs](#)

[Bridge-Based Active Domain Adaptation for Aspect Term Extraction](#)

[Zhuang Chen](#) | [Tieyun Qian](#)

[pdfbibabs](#)

[Multimodal Sentiment Detection Based on Multi-channel Graph Neural Networks](#)

[Xiaocui Yang](#) | [Shi Feng](#) | [Yifei Zhang](#) | [Daling Wang](#)

[pdfbibabs](#)

[Aspect-Category-Opinion-Sentiment Quadruple Extraction with Implicit Aspects and Opinions](#)

[Hongjie Cai](#) | [Rui Xia](#) | [Jianfei Yu](#)

[pdfbibabs](#)

[PASS: Perturb-and-Select Summarizer for Product Reviews](#)

[Nadav Oved](#) | [Ran Levy](#)

[pdfbibabs](#)

[Deep Differential Amplifier for Extractive Summarization](#)

[Ruipeng Jia](#) | [Yanan Cao](#) | [Fang Fang](#) | [Yuchen Zhou](#) | [Zheng Fang](#) | [Yanbing Liu](#) | [Shi Wang](#)

[pdfbibabs](#)

[Multi-TimeLine Summarization \(MTLS\): Improving Timeline Summarization by Generating Multiple Summaries](#)

[Yi Yu](#) | [Adam Jatowt](#) | [Antoine Doucet](#) | [Kazunari Sugiyama](#) | [Masatoshi Yoshikawa](#)

[pdfbibabs](#)

[Self-Supervised Multimodal Opinion Summarization](#)

[Jinbae Im](#) | [Moonki Kim](#) | [Hoyeop Lee](#) | [Hyunsouk Cho](#) | [Sehee Chung](#)

[pdfbibabs](#)

[A Training-free and Reference-free Summarization Evaluation Metric via Centrality-weighted Relevance and Self-referenced Redundancy](#)

[Wang Chen](#) | [Piji Li](#) | [Irwin King](#)

[pdfbibabs](#)

[DESCGEN: A Distantly Supervised Dataset for Generating Entity Descriptions](#)

[Weijia Shi](#) | [Mandar Joshi](#) | [Luke Zettlemoyer](#)

[pdfbibabs](#)

[Introducing Orthogonal Constraint in Structural Probes](#)

[Tomasz Limisiewicz](#) | [David Mareček](#)

[pdfbibabs](#)

[Hidden Killer: Invisible Textual Backdoor Attacks with Syntactic Trigger](#)

[Fanchao Qi](#) | [Mukai Li](#) | [Yangyi Chen](#) | [Zhengyan Zhang](#) | [Zhiyuan Liu](#) | [Yasheng Wang](#) | [Maosong Sun](#)

[pdfbibabs](#)

[Examining the Inductive Bias of Neural Language Models with Artificial Languages](#)

[Jennifer C. White](#) | [Ryan Cotterell](#)

[pdfbibabs](#)

[Explaining Contextualization in Language Models using Visual Analytics](#)

[Rita Sevastjanova](#) | [Aikaterini-Lida Kalouli](#) | [Christin Beck](#) | [Hanna Schäfer](#) | [Mennatallah El-Assady](#)

[pdfbibabs](#)

[Improving the Faithfulness of Attention-based Explanations with Task-specific Information for Text Classification](#)

[George Chrysostomou](#) | [Nikolaos Aletras](#)

[pdfbibabs](#)

[Generating Landmark Navigation Instructions from Maps as a Graph-to-Text Problem](#)

[Raphael Schumann](#) | [Stefan Riezler](#)

[pdfbibabs](#)

[E2E-VLP: End-to-End Vision-Language Pre-training Enhanced by Visual Learning](#)

[Haiyang Xu](#) | [Ming Yan](#) | [Chenliang Li](#) | [Bin Bi](#) | [Songfang Huang](#) | [Wenming Xiao](#) | [Fei Huang](#)

[pdfbibabs](#)

[Learning Relation Alignment for Calibrated Cross-modal Retrieval](#)

[Shuhuai Ren](#) | [Junyang Lin](#) | [Guangxiang Zhao](#) | [Rui Men](#) | [An Yang](#) | [Jingren Zhou](#) | [Xu Sun](#) | [Hongxia Yang](#)

[pdfbibabs](#)

[KM-BART: Knowledge Enhanced Multimodal BART for Visual Commonsense Generation](#)

[Yiran Xing](#) | [Zai Shi](#) | [Zhao Meng](#) | [Gerhard Lakemeyer](#) | [Yunpu Ma](#) | [Roger Wattenhofer](#)

[pdfbibabs](#)

[Cascaded Head-colliding Attention](#)

[Lin Zheng](#) | [Zhiyong Wu](#) | [Lingpeng Kong](#)

[pdfbibabs](#)

[Structural Knowledge Distillation: Tractably Distilling Information for Structured Predictor](#)

[Xinyu Wang](#) | [Yong Jiang](#) | [Zhaohui Yan](#) | [Zixia Jia](#) | [Nguyen Bach](#) | [Tao Wang](#) | [Zhongqiang Huang](#) | [Fei Huang](#) | [Kewei Tu](#)

[pdfbibabs](#)

[Parameter-efficient Multi-task Fine-tuning for Transformers via Shared Hypernetworks](#)

[Rabeeh Karimi Mahabadi](#) | [Sebastian Ruder](#) | [Mostafa Dehghani](#) | [James Henderson](#)

[pdfbibabs](#)

[COSY: COunterfactual SYntax for Cross-Lingual Understanding](#)

[Sicheng Yu](#) | [Hao Zhang](#) | [Yulei Niu](#) | [Qianru Sun](#) | [Jing Jiang](#)

[pdfbibabs](#)

[OoMMix: Out-of-manifold Regularization in Contextual Embedding Space for Text Classification](#)

[Seonghyeon Lee](#) | [Dongha Lee](#) | [Hwanjo Yu](#)

[pdfbibabs](#)

[Understanding and Countering Stereotypes: A Computational Approach to the Stereotype Content Model](#)

[Kathleen C. Fraser](#) | [Isar Nejadgholi](#) | [Svetlana Kiritchenko](#)

[pdfbibabs](#)

Structurizing Misinformation Stories via Rationalizing Fact-Checks

[Shan Jiang](#) | [Christo Wilson](#)

[pdfbibabs](#)

Modeling Language Usage and Listener Engagement in Podcasts

[Sravana Reddy](#) | [Mariya Lazarova](#) | [Yongze Yu](#) | [Rosie Jones](#)

[pdfbibabs](#)

Breaking Down the Invisible Wall of Informal Fallacies in Online Discussions

[Saumya Sahai](#) | [Oana Balalau](#) | [Roxana Horincar](#)

[pdfbibabs](#)

SocAoG: Incremental Graph Parsing for Social Relation Inference in Dialogues

[Liang Qiu](#) | [Yuan Liang](#) | [Yizhou Zhao](#) | [Pan Lu](#) | [Baolin Peng](#) | [Zhou Yu](#) | [Ying Nian Wu](#) | [Song-Chun Zhu](#)

[pdfbibabs](#)

TicketTalk: Toward human-level performance with end-to-end, transaction-based dialog systems

[Bill Byrne](#) | [Karthik Krishnamoorthi](#) | [Saravanan Ganesh](#) | [Mihir Kale](#)

[pdfbibabs](#)

Improving Dialog Systems for Negotiation with Personality Modeling

[Runzhe Yang](#) | [Jingxiao Chen](#) | [Karthik Narasimhan](#)

[pdfbibabs](#)

Learning from Perturbations: Diverse and Informative Dialogue Generation with Inverse Adversarial Training

[Wangchunshu Zhou](#) | [Qifei Li](#) | [Chenle Li](#)

[pdfbibabs](#)

Increasing Faithfulness in Knowledge-Grounded Dialogue with Controllable Features

[Hannah Rashkin](#) | [David Reitter](#) | [Gaurav Singh Tomar](#) | [Dipanjan Das](#)

[pdfbibabs](#)

CitationIE: Leveraging the Citation Graph for Scientific Information Extraction

[Vijay Viswanathan](#) | [Graham Neubig](#) | [Pengfei Liu](#)

[pdfbibabs](#)

From Discourse to Narrative: Knowledge Projection for Event Relation Extraction

[Jialong Tang](#) | [Hongyu Lin](#) | [Meng Liao](#) | [Yaojie Lu](#) | [Xianpei Han](#) | [Le Sun](#) | [Weijian Xie](#) | [Jin Xu](#)

[pdfbibabs](#)

AdvPicker: Effectively Leveraging Unlabeled Data via Adversarial Discriminator for Cross-Lingual NER

[Weile Chen](#) | [Huiqiang Jiang](#) | [Qianhui Wu](#) | [Börje Karlsson](#) | [Yi Guan](#)

[pdfbibabs](#)

Compare to The Knowledge: Graph Neural Fake News Detection with External Knowledge

[Linmei Hu](#) | [Tianchi Yang](#) | [Luhao Zhang](#) | [Wanjuan Zhong](#) | [Duyu Tang](#) | [Chuan Shi](#) | [Nan Duan](#) | [Ming Zhou](#)

[pdfbibabs](#)

Discontinuous Named Entity Recognition as Maximal Clique Discovery

[Yucheng Wang](#) | [Bowen Yu](#) | [Hongsong Zhu](#) | [Tingwen Liu](#) | [Nan Yu](#) | [Limin Sun](#)

[pdfbibabs](#)

LNN-EL: A Neuro-Symbolic Approach to Short-text Entity Linking

[Hang Jiang](#) | [Sairam Gurajada](#) | [Qiu hao Lu](#) | [Sumit Neelam](#) | [Lucian Popa](#) | [Prithviraj Sen](#) | [Yunyao Li](#) | [Alexander Gray](#)

[pdfbibabs](#)

Do Context-Aware Translation Models Pay the Right Attention?

[Kayo Yin](#) | [Patrick Fernandes](#) | [Danish Pruthi](#) | [Aditi Chaudhary](#) | [André F. T. Martins](#) | [Graham Neubig](#)

[pdfbibabs](#)

Adapting High-resource NMT Models to Translate Low-resource Related Languages without Parallel Data

[Wei-Jen Ko](#) | [Ahmed El-Kishky](#) | [Adithya Renduchintala](#) | [Vishrav Chaudhary](#) | [Naman Goyal](#) | [Francisco Guzmán](#) | [Pascale Fung](#) | [Philipp Koehn](#) | [Mona Diab](#)

[pdfbibabs](#)

Bilingual Lexicon Induction via Unsupervised Bitext Construction and Word Alignment

[Haoyue Shi](#) | [Luke Zettlemoyer](#) | [Sida I. Wang](#)

[pdfbibabs](#)

Multilingual Speech Translation from Efficient Finetuning of Pretrained Models

[Xian Li](#) | [Changhan Wang](#) | [Yun Tang](#) | [Chau Tran](#) | [Yuqing Tang](#) | [Juan Pino](#) | [Alexei Baevski](#) | [Alexis Conneau](#) | [Michael Auli](#)

[pdfbibabs](#)

Learning Faithful Representations of Causal Graphs

[Ananth Balashankar](#) | [Lakshminarayanan Subramanian](#)

[pdfbibabs](#)

What Context Features Can Transformer Language Models Use?

[Joe O'Connor](#) | [Jacob Andreas](#)

[pdfbibabs](#)

Integrated Directional Gradients: Feature Interaction Attribution for Neural NLP Models

[Sandipan Sikdar](#) | [Parantapa Bhattacharya](#) | [Kieran Heese](#)

[pdfbibabs](#)

DeCLUTR: Deep Contrastive Learning for Unsupervised Textual Representations

[John Giorgi](#) | [Osvald Nitski](#) | [Bo Wang](#) | [Gary Bader](#)

[pdfbibabs](#)

XLPT-AMR: Cross-Lingual Pre-Training via Multi-Task Learning for Zero-Shot AMR Parsing and Text Generation

[Dongqin Xu](#) | [Junhui Li](#) | [Muhua Zhu](#) | [Min Zhang](#) | [Guodong Zhou](#)

[pdfbibabs](#)

Span-based Semantic Parsing for Compositional Generalization

[Jonathan Herzig](#) | [Jonathan Berant](#)

[pdfbibabs](#)

Compositional Generalization and Natural Language Variation: Can a Semantic Parsing Approach Handle Both?

[Peter Shaw](#) | [Ming-Wei Chang](#) | [Panupong Pasupat](#) | [Kristina Toutanova](#)

[pdfbibabs](#)

A Targeted Assessment of Incremental Processing in Neural Language Models and Humans

[Ethan Wilcox](#) | [Pranali Vani](#) | [Roger Levy](#)

[pdfbibabs](#)

The Possible, the Plausible, and the Desirable: Event-Based Modality Detection for Language Processing

[Valentina Pyatkin](#) | [Shoval Sadde](#) | [Aynat Rubinstein](#) | [Paul Portner](#) | [Reut Tsarfaty](#)

[pdfbibabs](#)

To POS Tag or Not to POS Tag: The Impact of POS Tags on Morphological Learning in Low-Resource Settings

[Sarah Moeller](#) | [Ling Liu](#) | [Mans Hulden](#)

[pdfbibabs](#)

Prosodic segmentation for parsing spoken dialogue

[Elizabeth Nielsen](#) | [Mark Steedman](#) | [Sharon Goldwater](#)

[pdfbibabs](#)

VoxPopuli: A Large-Scale Multilingual Speech Corpus for Representation Learning, Semi-Supervised Learning and Interpretation

[Changhan Wang](#) | [Morgane Riviere](#) | [Ann Lee](#) | [Anne Wu](#) | [Chaitanya Talnikar](#) | [Daniel Haziza](#) | [Mary Williamson](#) | [Juan Pino](#) | [Emmanuel Dupoux](#)

[pdfbibabs](#)

Stereotyping Norwegian Salmon: An Inventory of Pitfalls in Fairness Benchmark Datasets

[Su Lin Blodgett](#) | [Gilsinia Lopez](#) | [Alexandra Olteanu](#) | [Robert Sim](#) | [Hanna Wallach](#)

[pdfbibabs](#)

Robust Knowledge Graph Completion with Stacked Convolutions and a Student Re-Ranking Network

[Justin Lovelace](#) | [Denis Newman-Griffis](#) | [Shikhar Vashishth](#) | [Jill Fain Lehman](#) | [Carolyn Rosé](#)

[pdfbibabs](#)

A DQN-based Approach to Finding Precise Evidences for Fact Verification

[Hai Wan](#) | [Haicheng Chen](#) | [Jianfeng Du](#) | [Weilin Luo](#) | [Rongzhen Ye](#)

[pdfbibabs](#)

The Art of Abstention: Selective Prediction and Error Regularization for Natural Language Processing

[Ji Xin](#) | [Raphael Tang](#) | [Yaoliang Yu](#) | [Jimmy Lin](#)

[pdfbibabs](#)

Unsupervised Out-of-Domain Detection via Pre-trained Transformers

[Keyang Xu](#) | [Tongzheng Ren](#) | [Shikun Zhang](#) | [Yihao Feng](#) | [Caiming Xiong](#)

[pdfbibabs](#)

MATE-KD: Masked Adversarial Text, a Companion to Knowledge Distillation

[Ahmad Rashid](#) | [Vasileios Lioutas](#) | [Mehdi Rezagholizadeh](#)

[pdfbibabs](#)

Selecting Informative Contexts Improves Language Model Fine-tuning

[Richard Antonello](#) | [Nicole Beckage](#) | [Javier Turek](#) | [Alexander Huth](#)

[pdfbibabs](#)

Explainable Prediction of Text Complexity: The Missing Preliminaries for Text Simplification

[Cristina Garbacea](#) | [Mengtian Guo](#) | [Samuel Carton](#) | [Qiaozhu Mei](#)

[pdfbibabs](#)

Multi-Task Retrieval for Knowledge-Intensive Tasks

[Jean Maillard](#) | [Vladimir Karpukhin](#) | [Fabio Petroni](#) | [Wen-tau Yih](#) | [Barlas Oguz](#) | [Veselin Stoyanov](#) | [Gargi Ghosh](#)

[pdfbibabs](#)

When Do You Need Billions of Words of Pretraining Data?

[Yian Zhang](#) | [Alex Warstadt](#) | [Xiaocheng Li](#) | [Samuel R. Bowman](#)

[pdfbibabs](#)

Analyzing the Source and Target Contributions to Predictions in Neural Machine Translation

[Elena Voita](#) | [Rico Sennrich](#) | [Ivan Titov](#)

[pdfbibabs](#)

Comparing Test Sets with Item Response Theory

[Clara Vania](#) | [Phu Mon Htut](#) | [William Huang](#) | [Dhara Mungra](#) | [Richard Yuanzhe Pang](#) | [Jason Phang](#) | [Haokun Liu](#) | [Kyunghyun Cho](#) | [Samuel R. Bowman](#)

[pdfbibabs](#)

Uncovering Constraint-Based Behavior in Neural Models via Targeted Fine-Tuning

[Forrest Davis](#) | [Marten van Schijndel](#)

[pdfbibabs](#)

More Identifiable yet Equally Performant Transformers for Text Classification

[Rishabh Bhardwaj](#) | [Navonil Majumder](#) | [Soujanya Poria](#) | [Eduard Hovy](#)

[pdfbibabs](#)

AugNLG: Few-shot Natural Language Generation using Self-trained Data Augmentation

[Xinnuo Xu](#) | [Guoyin Wang](#) | [Young-Bum Kim](#) | [Sungjin Lee](#)

[pdfbibabs](#)

Can vectors read minds better than experts? Comparing data augmentation strategies for the automated scoring of children's mindreading ability

[Venelin Kovatchev](#) | [Phillip Smith](#) | [Mark Lee](#) | [Rory Devine](#)

[pdfbibabs](#)

A Dataset and Baselines for Multilingual Reply Suggestion

[Mozhi Zhang](#) | [Wei Wang](#) | [Budhaditya Deb](#) | [Guoging Zheng](#) | [Milad Shokouhi](#) | [Ahmed Hassan Awadallah](#)

[pdfbibabs](#)

What Ingredients Make for an Effective Crowdsourcing Protocol for Difficult NLU Data Collection Tasks?

[Nikita Nangia](#) | [Saku Sugawara](#) | [Harsh Trivedi](#) | [Alex Warstadt](#) | [Clara Vania](#) | [Samuel R. Bowman](#)

[pdfbibabs](#)

Align Voting Behavior with Public Statements for Legislator Representation Learning

[Xinyi Mou](#) | [Zhongyu Wei](#) | [Lei Chen](#) | [Shangyi Ning](#) | [Yancheng He](#) | [Changjian Jiang](#) | [Xuanjing Huang](#)

[pdfbibabs](#)

Measure and Evaluation of Semantic Divergence across Two Languages

[Syrielle Montariol](#) | [Alexandre Allauzen](#)

[pdfbibabs](#)

Improving Zero-Shot Translation by Disentangling Positional Information

[Danni Liu](#) | [Jan Niehues](#) | [James Cross](#) | [Francisco Guzmán](#) | [Xian Li](#)

[pdfbibabs](#)

Common Sense Beyond English: Evaluating and Improving Multilingual Language Models for Commonsense Reasoning

[Bill Yuchen Lin](#) | [Seyeon Lee](#) | [Xiaoyang Qiao](#) | [Xiang Ren](#)

[pdfbibabs](#)

Attention Calibration for Transformer in Neural Machine Translation

[Yu Lu](#) | [Jiali Zeng](#) | [Jiajun Zhang](#) | [Shuangzhi Wu](#) | [Mu Li](#)

[pdfbibabs](#)

Diverse Pretrained Context Encodings Improve Document Translation

[Domenic Donato](#) | [Lei Yu](#) | [Chris Dyer](#)

[pdfbibabs](#)

Exploiting Language Relatedness for Low Web-Resource Language Model Adaptation: An Indic Languages Study

[Yash Khemchandani](#) | [Sarvesh Mehtani](#) | [Vaidehi Patil](#) | [Abhijeet Awasthi](#) | [Partha Talukdar](#) | [Sunita Sarawagi](#)

[pdfbibabs](#)

On Finding the K-best Non-projective Dependency Trees

[Ran Zmigrod](#) | [Tim Vieira](#) | [Ryan Cotterell](#)

[pdfbibabs](#)

Towards Argument Mining for Social Good: A Survey

[Eva Maria Vecchi](#) | [Neele Falk](#) | [Iman Jundi](#) | [Gabriella Lapesa](#)

[pdfbibabs](#)

Automated Generation of Storytelling Vocabulary from Photographs for use in AAC

[Mauricio Fontana de Vargas](#) | [Karyn Moffatt](#)

[pdfbibabs](#)

CLIP: A Dataset for Extracting Action Items for Physicians from Hospital Discharge Notes

[James Mullenbach](#) | [Yada Pruksachatkun](#) | [Sean Adler](#) | [Jennifer Seale](#) | [Jordan Swartz](#) | [Greg McKelvey](#) | [Hui Dai](#) | [Yi Yang](#) | [David Sontag](#)

[pdfbibabs](#)

Assessing Emoji Use in Modern Text Processing Tools

[Abu Awal Md Shoeb](#) | [Gerard de Melo](#)

[pdfbibabs](#)

Select, Extract and Generate: Neural Keyphrase Generation with Layer-wise Coverage Attention

[Wasi Ahmad](#) | [Xiao Bai](#) | [Soomin Lee](#) | [Kai-Wei Chang](#)

[pdfbibabs](#)

Factorising Meaning and Form for Intent-Preserving Paraphrasing

[Tom Hosking](#) | [Mirella Lapata](#)

[pdfbibabs](#)

AggGen: Ordering and Aggregating while Generating

[Xinnuo Xu](#) | [Ondřej Dušek](#) | [Verena Rieser](#) | [Ioannis Konstas](#)

[pdfbibabs](#)

Reflective Decoding: Beyond Unidirectional Generation with Off-the-Shelf Language Models

[Peter West](#) | [Ximing Lu](#) | [Ari Holtzman](#) | [Chandra Bhagavatula](#) | [Jena D. Hwang](#) | [Yejin Choi](#)

[pdfbibabs](#)

Towards Table-to-Text Generation with Numerical Reasoning

[Lya Hulliyyatus Suadaa](#) | [Hidetaka Kamigaito](#) | [Kotaro Funakoshi](#) | [Manabu Okumura](#) | [Hiroya Takamura](#)

[pdfbibabs](#)

BACO: A Background Knowledge- and Content-Based Framework for Citing Sentence Generation

[Yubin Ge](#) | [Ly Dinh](#) | [Xiaofeng Liu](#) | [Jinsong Su](#) | [Ziyao Lu](#) | [Ante Wang](#) | [Jana Diesner](#)

[pdfbibabs](#)

Language Model as an Annotator: Exploring DialoGPT for Dialogue Summarization

[Xiachong Feng](#) | [Xiaocheng Feng](#) | [Libo Qin](#) | [Bing Qin](#) | [Ting Liu](#)

[pdfbibabs](#)

Challenges in Information-Seeking QA: Unanswerable Questions and Paragraph Retrieval

[Akari Asai](#) | [Eunsol Choi](#)

[pdfbibabs](#)

A Gradually Soft Multi-Task and Data-Augmented Approach to Medical Question Understanding

[Khalil Mrini](#) | [Franck Dernoncourt](#) | [Seunghyun Yoon](#) | [Trung Bui](#) | [Walter Chang](#) | [Emilia Farcas](#) | [Ndapa Nakashole](#)

[pdfbibabs](#)

Leveraging Type Descriptions for Zero-shot Named Entity Recognition and Classification

[Rami Aly](#) | [Andreas Vlachos](#) | [Ryan McDonald](#)

[pdfbibabs](#)

MECT: Multi-Metadata Embedding based Cross-Transformer for Chinese Named Entity Recognition

[Shuang Wu](#) | [Xiaoning Song](#) | [Zhenhua Feng](#)

[pdfbibabs](#)

Factuality Assessment as Modal Dependency Parsing

[Jiarui Yao](#) | [Haoling Qiu](#) | [Jin Zhao](#) | [Bonan Min](#) | [Nianwen Xue](#)

[pdfbibabs](#)

Directed Acyclic Graph Network for Conversational Emotion Recognition

[Weizhou Shen](#) | [Siyue Wu](#) | [Yunyi Yang](#) | [Xiaojun Quan](#)

[pdfbibabs](#)

Improving Formality Style Transfer with Context-Aware Rule Injection

[Zonghai Yao](#) | [Hong Yu](#)

[pdfbibabs](#)

Topic-Driven and Knowledge-Aware Transformer for Dialogue Emotion Detection

[Lixing Zhu](#) | [Gabriele Pergola](#) | [Lin Gui](#) | [Deyu Zhou](#) | [Yulan He](#)

[pdfbibabs](#)

Syntopical Graphs for Computational Argumentation Tasks

[Joe Barrow](#) | [Rajiv Jain](#) | [Nedim Lipka](#) | [Franck Dernoncourt](#) | [Vlad Morariu](#) | [Varun Manjunatha](#) | [Douglas Oard](#) | [Philip Resnik](#) | [Henning Wachsmuth](#)

[pdfbibabs](#)

Stance Detection in COVID-19 Tweets

[Kyle Glandt](#) | [Sarthak Khanal](#) | [Yingjie Li](#) | [Doina Caragea](#) | [Cornelia Caragea](#)

[pdfbibabs](#)

Topic-Aware Evidence Reasoning and Stance-Aware Aggregation for Fact Verification

[Jiasheng Si](#) | [Deyu Zhou](#) | [Tongzhe Li](#) | [Xingyu Shi](#) | [Yulan He](#)

[pdfbibabs](#)

[Changes in European Solidarity Before and During COVID-19: Evidence from a Large Crowd- and Expert-Annotated Twitter Dataset](#)

[Alexandra Ils](#) | [Dan Liu](#) | [Daniela Grunow](#) | [Steffen Eger](#)

[pdfbibabs](#)

[Measuring Conversational Uptake: A Case Study on Student-Teacher Interactions](#)

[Dorottya Demszky](#) | [Jing Liu](#) | [Zid Mancenido](#) | [Julie Cohen](#) | [Heather Hill](#) | [Dan Jurafsky](#) | [Tatsunori Hashimoto](#)

[pdfbibabs](#)

[A Survey of Code-switching: Linguistic and Social Perspectives for Language Technologies](#)

[A. Seza Doğruöz](#) | [Sunayana Sitaram](#) | [Barbara E. Bullock](#) | [Almeida Jacqueline Toribio](#)

[pdfbibabs](#)

[Learning from the Worst: Dynamically Generated Datasets to Improve Online Hate Detection](#)

[Bertie Vidgen](#) | [Tristan Thrush](#) | [Zeerak Waseem](#) | [Douwe Kiela](#)

[pdfbibabs](#)

[InfoSurgeon: Cross-Media Fine-grained Information Consistency Checking for Fake News Detection](#)

[Yi Fung](#) | [Christopher Thomas](#) | [Revanth Gangi Reddy](#) | [Sandeep Polisetty](#) | [Heng Ji](#) | [Shih-Fu Chang](#) | [Kathleen McKeown](#) | [Mohit Bansal](#) | [Avi Sil](#)

[pdfbibabs](#)

[I like fish, especially dolphins: Addressing Contradictions in Dialogue Modeling](#)

[Yixin Nie](#) | [Mary Williamson](#) | [Mohit Bansal](#) | [Douwe Kiela](#) | [Jason Weston](#)

[pdfbibabs](#)

[A Sequence-to-Sequence Approach to Dialogue State Tracking](#)

[Yue Feng](#) | [Yang Wang](#) | [Hang Li](#)

[pdfbibabs](#)

[Discovering Dialog Structure Graph for Coherent Dialog Generation](#)

[Jun Xu](#) | [Zeyang Lei](#) | [Haifeng Wang](#) | [Zheng-Yu Niu](#) | [Hua Wu](#) | [Wanxiang Che](#)

[pdfbibabs](#)

[Dialogue Response Selection with Hierarchical Curriculum Learning](#)

[Yixuan Su](#) | [Deng Cai](#) | [Qingyu Zhou](#) | [Zibo Lin](#) | [Simon Baker](#) | [Yunbo Cao](#) | [Shuming Shi](#) | [Nigel Collier](#) | [Yan Wang](#)

[pdfbibabs](#)

[A Joint Model for Dropped Pronoun Recovery and Conversational Discourse Parsing in Chinese Conversational Speech](#)

[Jingxuan Yang](#) | [Kerui Xu](#) | [Jun Xu](#) | [Si Li](#) | [Sheng Gao](#) | [Jun Guo](#) | [Nianwen Xue](#) | [Ji-Rong Wen](#)

[pdfbibabs](#)

[A Systematic Investigation of KB-Text Embedding Alignment at Scale](#)

[Vardaan Pahuja](#) | [Yu Gu](#) | [Wenhu Chen](#) | [Mehdi Bahrami](#) | [Lei Liu](#) | [Wei-Peng Chen](#) | [Yu Su](#)

[pdfbibabs](#)

[Named Entity Recognition with Small Strongly Labeled and Large Weakly Labeled Data](#)

[Haoming Jiang](#) | [Danqing Zhang](#) | [Tianyu Cao](#) | [Bing Yin](#) | [Tuo Zhao](#)

[pdfbibabs](#)

[Ultra-Fine Entity Typing with Weak Supervision from a Masked Language Model](#)

[Hongliang Dai](#) | [Yangqiu Song](#) | [Haixun Wang](#)

[pdfbibabs](#)

Improving Named Entity Recognition by External Context Retrieving and Cooperative Learning

[Xinyu Wang](#) | [Yong Jiang](#) | [Nguyen Bach](#) | [Tao Wang](#) | [Zhongqiang Huang](#) | [Fei Huang](#) | [Kewei Tu](#)

[pdfbibabs](#)

Implicit Representations of Meaning in Neural Language Models

[Belinda Z. Li](#) | [Maxwell Nye](#) | [Jacob Andreas](#)

[pdfbibabs](#)

Causal Analysis of Syntactic Agreement Mechanisms in Neural Language Models

[Matthew Finlayson](#) | [Aaron Mueller](#) | [Sebastian Gehrmann](#) | [Stuart Shieber](#) | [Tal Linzen](#) | [Yonatan Belinkov](#)

[pdfbibabs](#)

Bird's Eye: Probing for Linguistic Graph Structures with a Simple Information-Theoretic Approach

[Yifan Hou](#) | [Mrinmaya Sachan](#)

[pdfbibabs](#)

Knowledgeable or Educated Guess? Revisiting Language Models as Knowledge Bases

[Boxi Cao](#) | [Hongyu Lin](#) | [Xianpei Han](#) | [Le Sun](#) | [Lingyong Yan](#) | [Meng Liao](#) | [Tong Xue](#) | [Jin Xu](#)

[pdfbibabs](#)

Poisoning Knowledge Graph Embeddings via Relation Inference Patterns

[Peru Bhardwaj](#) | [John Kelleher](#) | [Luca Costabello](#) | [Declan O'Sullivan](#)

[pdfbibabs](#)

Bad Seeds: Evaluating Lexical Methods for Bias Measurement

[Maria Antoniak](#) | [David Mimno](#)

[pdfbibabs](#)

A Survey of Race, Racism, and Anti-Racism in NLP

[Anjalie Field](#) | [Su Lin Blodgett](#) | [Zeerak Waseem](#) | [Yulia Tsvetkov](#)

[pdfbibabs](#)

Intrinsic Bias Metrics Do Not Correlate with Application Bias

[Seraphina Goldfarb-Tarrant](#) | [Rebecca Marchant](#) | [Ricardo Muñoz Sánchez](#) | [Mugdha Pandya](#) | [Adam Lopez](#)

[pdfbibabs](#)

RedditBias: A Real-World Resource for Bias Evaluation and Debiasing of Conversational Language Models

[Soumya Barikeri](#) | [Anne Lauscher](#) | [Ivan Vulić](#) | [Goran Glavaš](#)

[pdfbibabs](#)

Contributions of Transformer Attention Heads in Multi- and Cross-lingual Tasks

[Weicheng Ma](#) | [Kai Zhang](#) | [Renze Lou](#) | [Lili Wang](#) | [Soroush Vosoughi](#)

[pdfbibabs](#)

Crafting Adversarial Examples for Neural Machine Translation

[Xinze Zhang](#) | [Junzhe Zhang](#) | [Zhenhua Chen](#) | [Kun He](#)

[pdfbibabs](#)

UXLA: A Robust Unsupervised Data Augmentation Framework for Zero-Resource Cross-Lingual NLP

[M Saiful Bari](#) | [Tasnim Mohiuddin](#) | [Shafiq Joty](#)

[pdfbibabs](#)

[Glancing Transformer for Non-Autoregressive Neural Machine Translation](#)

[Lihua Qian](#) | [Hao Zhou](#) | [Yu Bao](#) | [Mingxuan Wang](#) | [Lin Qiu](#) | [Weinan Zhang](#) | [Yong Yu](#) | [Lei Li](#)

[pdfbibabs](#)

[Hierarchical Context-aware Network for Dense Video Event Captioning](#)

[Lei Ji](#) | [Xianglin Guo](#) | [Haoyang Huang](#) | [Xilin Chen](#)

[pdfbibabs](#)

[Control Image Captioning Spatially and Temporally](#)

[Kun Yan](#) | [Lei Ji](#) | [Huaishao Luo](#) | [Ming Zhou](#) | [Nan Duan](#) | [Shuai Ma](#)

[pdfbibabs](#)

[Edited Media Understanding Frames: Reasoning About the Intent and Implications of Visual Misinformation](#)

[Jeff Da](#) | [Maxwell Forbes](#) | [Rowan Zellers](#) | [Anthony Zheng](#) | [Jena D. Hwang](#) | [Antoine Bosselut](#) | [Yejin Choi](#)

[pdfbibabs](#)

[PIGLeT: Language Grounding Through Neuro-Symbolic Interaction in a 3D World](#)

[Rowan Zellers](#) | [Ari Holtzman](#) | [Matthew Peters](#) | [Roosbeh Mottaghi](#) | [Aniruddha Kembhavi](#) | [Ali Farhadi](#) | [Yejin Choi](#)

[pdfbibabs](#)

[Modeling Fine-Grained Entity Types with Box Embeddings](#)

[Yasumasa Onoe](#) | [Michael Boratko](#) | [Andrew McCallum](#) | [Greg Durrett](#)

[pdfbibabs](#)

[ChineseBERT: Chinese Pretraining Enhanced by Glyph and Pinyin Information](#)

[Zijun Sun](#) | [Xiaoya Li](#) | [Xiaofei Sun](#) | [Yuxian Meng](#) | [Xiang Ao](#) | [Qing He](#) | [Fei Wu](#) | [Jiwei Li](#)

[pdfbibabs](#)

[Weight Distillation: Transferring the Knowledge in Neural Network Parameters](#)

[Ye Lin](#) | [Yanyang Li](#) | [Ziyang Wang](#) | [Bei Li](#) | [Quan Du](#) | [Tong Xiao](#) | [Jingbo Zhu](#)

[pdfbibabs](#)

[Optimizing Deeper Transformers on Small Datasets](#)

[Peng Xu](#) | [Dhruv Kumar](#) | [Wei Yang](#) | [Wenjie Zi](#) | [Keyi Tang](#) | [Chenyang Huang](#) | [Jackie Chi Kit Cheung](#) | [Simon J.D. Prince](#) | [Yanshuai Cao](#)

[pdfbibabs](#)

[BERTAC: Enhancing Transformer-based Language Models with Adversarially Pretrained Convolutional Neural Networks](#)

[Jong-Hoon Oh](#) | [Ryu Iida](#) | [Julien Klotzer](#) | [Kentaro Torisawa](#)

[pdfbibabs](#)

[COVID-Fact: Fact Extraction and Verification of Real-World Claims on COVID-19 Pandemic](#)

[Arkadiy Saakyan](#) | [Tuhin Chakrabarty](#) | [Smaranda Muresan](#)

[pdfbibabs](#)

[Explaining Relationships Between Scientific Documents](#)

[Kelvin Luu](#) | [Xinyi Wu](#) | [Rik Koncel-Kedziorski](#) | [Kyle Lo](#) | [Isabel Cachola](#) | [Noah A. Smith](#)

[pdfbibabs](#)

[IrEne: Interpretable Energy Prediction for Transformers](#)

[Qingqing Cao](#) | [Yash Kumar Lal](#) | [Harsh Trivedi](#) | [Aruna Balasubramanian](#) | [Niranjan Balasubramanian](#)

[pdfbibabs](#)

Mitigating Bias in Session-based Cyberbullying Detection: A Non-Compromising Approach

[Lu Cheng](#) | [Ahmadreza Mosallanezhad](#) | [Yasin Silva](#) | [Deborah Hall](#) | [Huan Liu](#)

[pdfbibabs](#)

PlotCoder: Hierarchical Decoding for Synthesizing Visualization Code in Programmatic Context

[Xinyun Chen](#) | [Linyuan Gong](#) | [Alvin Cheung](#) | [Dawn Song](#)

[pdfbibabs](#)

Changing the World by Changing the Data

[Anna Rogers](#)

[pdfbibabs](#)

EarlyBERT: Efficient BERT Training via Early-bird Lottery Tickets

[Xiaohan Chen](#) | [Yu Cheng](#) | [Shuohang Wang](#) | [Zhe Gan](#) | [Zhangyang Wang](#) | [Jingjing Liu](#)

[pdfbibabs](#)

On the Effectiveness of Adapter-based Tuning for Pretrained Language Model Adaptation

[Ruidan He](#) | [Linlin Liu](#) | [Hai Ye](#) | [Qingyu Tan](#) | [Bosheng Ding](#) | [Liyang Cheng](#) | [Jiawei Low](#) | [Lidong Bing](#) | [Luo Si](#)

[pdfbibabs](#)

Data Augmentation for Text Generation Without Any Augmented Data

[Wei Bi](#) | [Huayang Li](#) | [Jiacheng Huang](#)

[pdfbibabs](#)

Integrating Semantics and Neighborhood Information with Graph-Driven Generative Models for Document Retrieval

[Zijing Ou](#) | [Qinliang Su](#) | [Jianxing Yu](#) | [Bang Liu](#) | [Jingwen Wang](#) | [Ruihui Zhao](#) | [Changyou Chen](#) | [Yefeng Zheng](#)

[pdfbibabs](#)

SMURF: SeMantic and linguistic UnderStanding Fusion for Caption Evaluation via Typicality Analysis

[Joshua Feinglass](#) | [Yezhou Yang](#)

[pdfbibabs](#)

KaggleDBQA: Realistic Evaluation of Text-to-SQL Parsers

[Chia-Hsuan Lee](#) | [Oleksandr Polozov](#) | [Matthew Richardson](#)

[pdfbibabs](#)

QASR: QCRI Aljazeera Speech Resource A Large Scale Annotated Arabic Speech Corpus

[Hamdy Mubarak](#) | [Amir Hussein](#) | [Shammur Absar Chowdhury](#) | [Ahmed Ali](#)

[pdfbibabs](#)

An Empirical Study on Hyperparameter Optimization for Fine-Tuning Pre-trained Language Models

[Xueqing Liu](#) | [Chi Wang](#)

[pdfbibabs](#)

Better than Average: Paired Evaluation of NLP systems

[Maxime Peyrard](#) | [Wei Zhao](#) | [Steffen Eger](#) | [Robert West](#)

[pdfbibabs](#)

Chase: A Large-Scale and Pragmatic Chinese Dataset for Cross-Database Context-Dependent Text-to-SQL

[Jiaqi Guo](#) | [Ziliang Si](#) | [Yu Wang](#) | [Qian Liu](#) | [Ming Fan](#) | [Jian-Guang Lou](#) | [Zijiang Yang](#) | [Ting Liu](#)

[pdfbibabs](#)

CLINE: Contrastive Learning with Semantic Negative Examples for Natural Language Understanding

[Dong Wang](#) | [Ning Ding](#) | [Piji Li](#) | [Haitao Zheng](#)

[pdfbibabs](#)

Tree-Structured Topic Modeling with Nonparametric Neural Variational Inference

[Ziye Chen](#) | [Cheng Ding](#) | [Zusheng Zhang](#) | [Yanghui Rao](#) | [Haoran Xie](#)

[pdfbibabs](#)

ExCAR: Event Graph Knowledge Enhanced Explainable Causal Reasoning

[Li Du](#) | [Xiao Ding](#) | [Kai Xiong](#) | [Ting Liu](#) | [Bing Qin](#)

[pdfbibabs](#)

Distributed Representations of Emotion Categories in Emotion Space

[Xiangyu Wang](#) | [Chengqing Zong](#)

[pdfbibabs](#)

Style is NOT a single variable: Case Studies for Cross-Stylistic Language Understanding

[Dongyeop Kang](#) | [Eduard Hovy](#)

[pdfbibabs](#)

DynaSent: A Dynamic Benchmark for Sentiment Analysis

[Christopher Potts](#) | [Zhengxuan Wu](#) | [Atticus Geiger](#) | [Douwe Kiela](#)

[pdfbibabs](#)

A Hierarchical VAE for Calibrating Attributes while Generating Text using Normalizing Flow

[Bidisha Samanta](#) | [Mohit Agrawal](#) | [Niloy Ganguly](#)

[pdfbibabs](#)

A Unified Generative Framework for Aspect-based Sentiment Analysis

[Hang Yan](#) | [Junqi Dai](#) | [Tuo Ji](#) | [Xipeng Qiu](#) | [Zheng Zhang](#)

[pdfbibabs](#)

Discovering Dialogue Slots with Weak Supervision

[Vojtěch Hudeček](#) | [Ondřej Dušek](#) | [Zhou Yu](#)

[pdfbibabs](#)

Enhancing the generalization for Intent Classification and Out-of-Domain Detection in SLU

[Yilin Shen](#) | [Yen-Chang Hsu](#) | [Avik Ray](#) | [Hongxia Jin](#)

[pdfbibabs](#)

ProtAugment: Intent Detection Meta-Learning through Unsupervised Diverse Paraphrasing

[Thomas Dopierre](#) | [Christophe Gravier](#) | [Wilfried Logerais](#)

[pdfbibabs](#)

Robustness Testing of Language Understanding in Task-Oriented Dialog

[Jiexi Liu](#) | [Ryuichi Takanobu](#) | [Jiaxin Wen](#) | [Dazhen Wan](#) | [Hongguang Li](#) | [Weiran Nie](#) | [Cheng Li](#) | [Wei Peng](#) | [Minlie Huang](#)

[pdfbibabs](#)

Comprehensive Study: How the Context Information of Different Granularity Affects Dialogue State Tracking?

[Puhai Yang](#) | [Heyan Huang](#) | [Xian-Ling Mao](#)

[pdfbibabs](#)

OTters: One-turn Topic Transitions for Open-Domain Dialogue

[Karin Sevegnani](#) | [David M. Howcroft](#) | [Ioannis Konstas](#) | [Verena Rieser](#)

[pdfbibabs](#)

Towards Robustness of Text-to-SQL Models against Synonym Substitution

[Yujian Gan](#) | [Xinyun Chen](#) | [Qiuping Huang](#) | [Matthew Purver](#) | [John R. Woodward](#) | [Jinxia Xie](#) | [Pengsheng Huang](#)

[pdfbibabs](#)

KACE: Generating Knowledge Aware Contrastive Explanations for Natural Language Inference

[Qianglong Chen](#) | [Feng Ji](#) | [Xiangji Zeng](#) | [Feng-Lin Li](#) | [Ji Zhang](#) | [Haiqing Chen](#) | [Yin Zhang](#)

[pdfbibabs](#)

Self-Guided Contrastive Learning for BERT Sentence Representations

[Taeuk Kim](#) | [Kang Min Yoo](#) | [Sang-goo Lee](#)

[pdfbibabs](#)

LGESQL: Line Graph Enhanced Text-to-SQL Model with Mixed Local and Non-Local Relations

[Ruisheng Cao](#) | [Lu Chen](#) | [Zhi Chen](#) | [Yanbin Zhao](#) | [Su Zhu](#) | [Kai Yu](#)

[pdfbibabs](#)

Multi-stage Pre-training over Simplified Multimodal Pre-training Models

[Tongtong Liu](#) | [Fangxiang Feng](#) | [Xiaojie Wang](#)

[pdfbibabs](#)

Beyond Sentence-Level End-to-End Speech Translation: Context Helps

[Biao Zhang](#) | [Ivan Titov](#) | [Barry Haddow](#) | [Rico Sennrich](#)

[pdfbibabs](#)

LayoutLMv2: Multi-modal Pre-training for Visually-rich Document Understanding

[Yang Xu](#) | [Yiheng Xu](#) | [Tengchao Lv](#) | [Lei Cui](#) | [Furu Wei](#) | [Guoxin Wang](#) | [Yijuan Lu](#) | [Dinei Florencio](#) | [Cha Zhang](#) | [Wanxiang Che](#) | [Min Zhang](#) | [Lidong Zhou](#)

[pdfbibabs](#)

UNIMO: Towards Unified-Modal Understanding and Generation via Cross-Modal Contrastive Learning

[Wei Li](#) | [Can Gao](#) | [Guocheng Niu](#) | [Xinyan Xiao](#) | [Hao Liu](#) | [Jiachen Liu](#) | [Hua Wu](#) | [Haifeng Wang](#)

[pdfbibabs](#)

Missing Modality Imagination Network for Emotion Recognition with Uncertain Missing Modalities

[Jinming Zhao](#) | [Ruichen Li](#) | [Qin Jin](#)

[pdfbibabs](#)

Stacked Acoustic-and-Textual Encoding: Integrating the Pre-trained Models into Speech Translation Encoders

[Chen Xu](#) | [Bojie Hu](#) | [Yanyang Li](#) | [Yuhao Zhang](#) | [Shen Huang](#) | [Qi Ju](#) | [Tong Xiao](#) | [Jingbo Zhu](#)

[pdfbibabs](#)

N-ary Constituent Tree Parsing with Recursive Semi-Markov Model

[Xin Xin](#) | [Jinlong Li](#) | [Zeqi Tan](#)

[pdfbibabs](#)

Automated Concatenation of Embeddings for Structured Prediction

[Xinyu Wang](#) | [Yong Jiang](#) | [Nguyen Bach](#) | [Tao Wang](#) | [Zhongqiang Huang](#) | [Fei Huang](#) | [Kewei Tu](#)

[pdfbibabs](#)

Multi-View Cross-Lingual Structured Prediction with Minimum Supervision

[Zechuan Hu](#) | [Yong Jiang](#) | [Nguyen Bach](#) | [Tao Wang](#) | [Zhongqiang Huang](#) | [Fei Huang](#) | [Kewei Tu](#)

[pdfbibabs](#)

[The Limitations of Limited Context for Constituency Parsing](#)

[Yuchen Li](#) | [Andrej Risteski](#)

[pdfbibabs](#)

[Neural Bi-Lexicalized PCFG Induction](#)

[Songlin Yang](#) | [Yanpeng Zhao](#) | [Kewei Tu](#)

[pdfbibabs](#)

[Ruddit: Norms of Offensiveness for English Reddit Comments](#)

[Rishav Hada](#) | [Sohi Sudhir](#) | [Pushkar Mishra](#) | [Helen Yannakoudakis](#) | [Saif M. Mohammad](#) | [Ekaterina Shutova](#)

[pdfbibabs](#)

[Towards Quantifiable Dialogue Coherence Evaluation](#)

[Zheng Ye](#) | [Liucun Lu](#) | [Lishan Huang](#) | [Liang Lin](#) | [Xiaodan Liang](#)

[pdfbibabs](#)

[Assessing the Representations of Idiomaticity in Vector Models with a Noun Compound Dataset Labeled at Type and Token Levels](#)

[Marcos Garcia](#) | [Tiago Kramer Vieira](#) | [Carolina Scarton](#) | [Marco Idiart](#) | [Aline Villavicencio](#)

[pdfbibabs](#)

[Factoring Statutory Reasoning as Language Understanding Challenges](#)

[Nils Holzenberger](#) | [Benjamin Van Durme](#)

[pdfbibabs](#)

[Evaluating Evaluation Measures for Ordinal Classification and Ordinal Quantification](#)

[Tetsuya Sakai](#)

[pdfbibabs](#)

[Interpretable and Low-Resource Entity Matching via Decoupling Feature Learning from Decision Making](#)

[Zijun Yao](#) | [Chengjiang Li](#) | [Tiansi Dong](#) | [Xin Lv](#) | [Jifan Yu](#) | [Lei Hou](#) | [Juanzi Li](#) | [Yichi Zhang](#) | [Zelin Dai](#)

[pdfbibabs](#)

[Locate and Label: A Two-stage Identifier for Nested Named Entity Recognition](#)

[Yongliang Shen](#) | [Xinyin Ma](#) | [Zeqi Tan](#) | [Shuai Zhang](#) | [Wen Wang](#) | [Weiming Lu](#)

[pdfbibabs](#)

[Text2Event: Controllable Sequence-to-Structure Generation for End-to-end Event Extraction](#)

[Yaojie Lu](#) | [Hongyu Lin](#) | [Jin Xu](#) | [Xianpei Han](#) | [Jialong Tang](#) | [Annan Li](#) | [Le Sun](#) | [Meng Liao](#) | [Shaoyi Chen](#)

[pdfbibabs](#)

[A Large-Scale Chinese Multimodal NER Dataset with Speech Clues](#)

[Dianbo Sui](#) | [Zhengkun Tian](#) | [Yubo Chen](#) | [Kang Liu](#) | [Jun Zhao](#)

[pdfbibabs](#)

[A Neural Transition-based Joint Model for Disease Named Entity Recognition and Normalization](#)

[Zongcheng Ji](#) | [Tian Xia](#) | [Mei Han](#) | [Jing Xiao](#)

[pdfbibabs](#)

[OntoED: Low-resource Event Detection with Ontology Embedding](#)

[Shumin Deng](#) | [Ningyu Zhang](#) | [Luoqi Li](#) | [Chen Hui](#) | [Tou Huaixiao](#) | [Mosha Chen](#) | [Fei Huang](#) | [Huajun Chen](#)

[pdfbibabs](#)

Self-Training Sampling with Monolingual Data Uncertainty for Neural Machine Translation

[Wenxiang Jiao](#) | [Xing Wang](#) | [Zhaopeng Tu](#) | [Shuming Shi](#) | [Michael Lyu](#) | [Irwin King](#)

[pdfbibabs](#)

Breaking the Corpus Bottleneck for Context-Aware Neural Machine Translation with Cross-Task Pre-training

[Linqing Chen](#) | [Junhui Li](#) | [Zhengxian Gong](#) | [Boxing Chen](#) | [Weihua Luo](#) | [Min Zhang](#) | [Guodong Zhou](#)

[pdfbibabs](#)

Guiding Teacher Forcing with Seer Forcing for Neural Machine Translation

[Yang Feng](#) | [Shuhao Gu](#) | [Dengji Guo](#) | [Zhengxin Yang](#) | [Chenze Shao](#)

[pdfbibabs](#)

Cascade versus Direct Speech Translation: Do the Differences Still Make a Difference?

[Luisa Bentivogli](#) | [Mauro Cettolo](#) | [Marco Gaido](#) | [Alina Karakanta](#) | [Alberto Martinelli](#) | [Matteo Negri](#) | [Marco Turchi](#)

[pdfbibabs](#)

Unsupervised Neural Machine Translation for Low-Resource Domains via Meta-Learning

[Cheonbok Park](#) | [Yunwon Tae](#) | [TaeHee Kim](#) | [Soyoung Yang](#) | [Mohammad Azam Khan](#) | [Lucy Park](#) | [Jaegul Choo](#)

[pdfbibabs](#)

Lightweight Cross-Lingual Sentence Representation Learning

[Zhuoyuan Mao](#) | [Prakhar Gupta](#) | [Chenhui Chu](#) | [Martin Jaggi](#) | [Sadao Kurohashi](#)

[pdfbibabs](#)

ERNIE-Doc: A Retrospective Long-Document Modeling Transformer

[SiYu Ding](#) | [Junyuan Shang](#) | [Shuohuan Wang](#) | [Yu Sun](#) | [Hao Tian](#) | [Hua Wu](#) | [Haifeng Wang](#)

[pdfbibabs](#)

Marginal Utility Diminishes: Exploring the Minimum Knowledge for BERT Knowledge Distillation

[Yuanxin Liu](#) | [Fandong Meng](#) | [Zheng Lin](#) | [Weiping Wang](#) | [Jie Zhou](#)

[pdfbibabs](#)

Rational LAMOL: A Rationale-based Lifelong Learning Framework

[Kasidis Kanwatchara](#) | [Thanapapas Horsuwan](#) | [Piyawat Lertvittayakumjorn](#) | [Boonserm Kijisirikul](#) | [Peerapon Vateekul](#)

[pdfbibabs](#)

EnSLM: Ensemble Language Model for Data Diversity by Semantic Clustering

[Zhibin Duan](#) | [Hao Zhang](#) | [Chaojie Wang](#) | [Zhengjue Wang](#) | [Bo Chen](#) | [Mingyuan Zhou](#)

[pdfbibabs](#)

LeeBERT: Learned Early Exit for BERT with cross-level optimization

[Wei Zhu](#)

[pdfbibabs](#)

Unsupervised Extractive Summarization-Based Representations for Accurate and Explainable Collaborative Filtering

[Reinald Adrian Pugoy](#) | [Hung-Yu Kao](#)

[pdfbibabs](#)

PLOME: Pre-training with Misspelled Knowledge for Chinese Spelling Correction

[Shulin Liu](#) | [Tao Yang](#) | [Tianchi Yue](#) | [Feng Zhang](#) | [Di Wang](#)

[pdfbibabs](#)

Competence-based Multimodal Curriculum Learning for Medical Report Generation

[Fenglin Liu](#) | [Shen Ge](#) | [Xian Wu](#)

[pdfbibabs](#)

Learning Syntactic Dense Embedding with Correlation Graph for Automatic Readability Assessment

[Xinying Qiu](#) | [Yuan Chen](#) | [Hanwu Chen](#) | [Jian-Yun Nie](#) | [Yuming Shen](#) | [Dawei Lu](#)

[pdfbibabs](#)

Meta-KD: A Meta Knowledge Distillation Framework for Language Model Compression across Domains

[Haojie Pan](#) | [Chengyu Wang](#) | [Minghui Qiu](#) | [Yichang Zhang](#) | [Yaliang Li](#) | [Jun Huang](#)

[pdfbibabs](#)

A Semantic-based Method for Unsupervised Commonsense Question Answering

[Yilin Niu](#) | [Fei Huang](#) | [Jiaming Liang](#) | [Wenkai Chen](#) | [Xiaoyan Zhu](#) | [Minlie Huang](#)

[pdfbibabs](#)

Explanations for CommonsenseQA: New Dataset and Models

[Shourya Aggarwal](#) | [Divyanshu Mandowara](#) | [Vishwajeet Agrawal](#) | [Dinesh Khandelwal](#) | [Parag Singla](#) | [Dinesh Garg](#)

[pdfbibabs](#)

Few-Shot Question Answering by Pretraining Span Selection

[Ori Ram](#) | [Yuval Kirstain](#) | [Jonathan Berant](#) | [Amir Globerson](#) | [Omer Levy](#)

[pdfbibabs](#)

UnitedQA: A Hybrid Approach for Open Domain Question Answering

[Hao Cheng](#) | [Yelong Shen](#) | [Xiaodong Liu](#) | [Pengcheng He](#) | [Weizhu Chen](#) | [Jianfeng Gao](#)

[pdfbibabs](#)

Database reasoning over text

[James Thorne](#) | [Majid Yazdani](#) | [Marzieh Saeidi](#) | [Fabrizio Silvestri](#) | [Sebastian Riedel](#) | [Alon Halevy](#)

[pdfbibabs](#)

Online Learning Meets Machine Translation Evaluation: Finding the Best Systems with the Least Human Effort

[Vânia Mendonça](#) | [Ricardo Rei](#) | [Luisa Coheur](#) | [Alberto Sardinha](#) | [Ana Lúcia Santos](#)

[pdfbibabs](#)

How Good is Your Tokenizer? On the Monolingual Performance of Multilingual Language Models

[Phillip Rust](#) | [Jonas Pfeiffer](#) | [Ivan Vulić](#) | [Sebastian Ruder](#) | [Iryna Gurevych](#)

[pdfbibabs](#)

Evaluating morphological typology in zero-shot cross-lingual transfer

[Antonio Martínez-García](#) | [Toni Badia](#) | [Jeremy Barnes](#)

[pdfbibabs](#)

From Machine Translation to Code-Switching: Generating High-Quality Code-Switched Text

[Ishan Tarunesh](#) | [Syamantak Kumar](#) | [Preethi Jyothi](#)

[pdfbibabs](#)

Fast and Accurate Neural Machine Translation with Translation Memory

[Qiuxiang He](#) | [Guoping Huang](#) | [Qu Cui](#) | [Li Li](#) | [Lemao Liu](#)

[pdfbibabs](#)

Annotating Online Misogyny

[Philine Zeinert](#) | [Nanna Inie](#) | [Leon Derczynski](#)

[pdfbibabs](#)

Few-NERD: A Few-shot Named Entity Recognition Dataset

[Ning Ding](#) | [Guangwei Xu](#) | [Yulin Chen](#) | [Xiaobin Wang](#) | [Xu Han](#) | [Pengjun Xie](#) | [Haitao Zheng](#) | [Zhiyuan Liu](#)

[pdfbibabs](#)

MultiMET: A Multimodal Dataset for Metaphor Understanding

[Dongyu Zhang](#) | [Minghao Zhang](#) | [Heting Zhang](#) | [Liang Yang](#) | [Hongfei Lin](#)

[pdfbibabs](#)

Human-in-the-Loop for Data Collection: a Multi-Target Counter Narrative Dataset to Fight Online Hate Speech

[Margherita Fanton](#) | [Helena Bonaldi](#) | [Serra Sinem Tekiroğlu](#) | [Marco Guerini](#)

[pdfbibabs](#)

Can Generative Pre-trained Language Models Serve As Knowledge Bases for Closed-book QA?

[Cunxiang Wang](#) | [Pai Liu](#) | [Yue Zhang](#)

[pdfbibabs](#)

Joint Models for Answer Verification in Question Answering Systems

[Zeyu Zhang](#) | [Thuy Vu](#) | [Alessandro Moschitti](#)

[pdfbibabs](#)

Answering Ambiguous Questions through Generative Evidence Fusion and Round-Trip Prediction

[Yifan Gao](#) | [Henghui Zhu](#) | [Patrick Ng](#) | [Cicero Nogueira dos Santos](#) | [Zhiguo Wang](#) | [Feng Nan](#) | [Dejiao Zhang](#) | [Ramesh Nallapati](#) | [Andrew O. Arnold](#) | [Bing Xiang](#)

[pdfbibabs](#)

TAT-QA: A Question Answering Benchmark on a Hybrid of Tabular and Textual Content in Finance

[Fengbin Zhu](#) | [Wenqiang Lei](#) | [Youcheng Huang](#) | [Chao Wang](#) | [Shuo Zhang](#) | [Jiancheng Lv](#) | [Fuli Feng](#) | [Tat-Seng Chua](#)

[pdfbibabs](#)

Modeling Transitions of Focal Entities for Conversational Knowledge Base Question Answering

[Yunshi Lan](#) | [Jing Jiang](#)

[pdfbibabs](#)

Evidence-based Factual Error Correction

[James Thorne](#) | [Andreas Vlachos](#)

[pdfbibabs](#)

Probabilistic, Structure-Aware Algorithms for Improved Variety, Accuracy, and Coverage of AMR Alignments

[Austin Blodgett](#) | [Nathan Schneider](#)

[pdfbibabs](#)

Meta-Learning to Compositionally Generalize

[Henry Conklin](#) | [Bailin Wang](#) | [Kenny Smith](#) | [Ivan Titov](#)

[pdfbibabs](#)

Taming Pre-trained Language Models with N-gram Representations for Low-Resource Domain Adaptation

[Shizhe Diao](#) | [Ruijia Xu](#) | [Hongjin Su](#) | [Yilei Jiang](#) | [Yan Song](#) | [Tong Zhang](#)

[pdfbibabs](#)

ERICA: Improving Entity and Relation Understanding for Pre-trained Language Models via Contrastive Learning

[Yujia Qin](#) | [Yankai Lin](#) | [Ryuichi Takanobu](#) | [Zhiyuan Liu](#) | [Peng Li](#) | [Heng Ji](#) | [Minlie Huang](#) | [Maosong Sun](#) | [Jie Zhou](#)

[pdfbibabs](#)

Position Bias Mitigation: A Knowledge-Aware Graph Model for Emotion Cause Extraction

[Hanqi Yan](#) | [Lin Gui](#) | [Gabriele Pergola](#) | [Yulan He](#)

[pdfbibabs](#)

Every Bite Is an Experience: Key Point Analysis of Business Reviews

[Roy Bar-Haim](#) | [Lilach Eden](#) | [Yoav Kantor](#) | [Roni Friedman](#) | [Noam Slonim](#)

[pdfbibabs](#)

Structured Sentiment Analysis as Dependency Graph Parsing

[Jeremy Barnes](#) | [Robin Kurtz](#) | [Stephan Oepen](#) | [Lilja Øvrelid](#) | [Erik Velldal](#)

[pdfbibabs](#)

Consistency Regularization for Cross-Lingual Fine-Tuning

[Bo Zheng](#) | [Li Dong](#) | [Shaohan Huang](#) | [Wenhui Wang](#) | [Zewen Chi](#) | [Saksham Singhal](#) | [Wanxiang Che](#) | [Ting Liu](#) | [Xia Song](#) | [Furu Wei](#)

[pdfbibabs](#)

Improving Pretrained Cross-Lingual Language Models via Self-Labeled Word Alignment

[Zewen Chi](#) | [Li Dong](#) | [Bo Zheng](#) | [Shaohan Huang](#) | [Xian-Ling Mao](#) | [Heyan Huang](#) | [Furu Wei](#)

[pdfbibabs](#)

Rejuvenating Low-Frequency Words: Making the Most of Parallel Data in Non-Autoregressive Translation

[Liang Ding](#) | [Longyue Wang](#) | [Xuebo Liu](#) | [Derek F. Wong](#) | [Dacheng Tao](#) | [Zhaopeng Tu](#)

[pdfbibabs](#)

G-Transformer for Document-Level Machine Translation

[Guangsheng Bao](#) | [Yue Zhang](#) | [Zhiyang Teng](#) | [Boxing Chen](#) | [Weihua Luo](#)

[pdfbibabs](#)

Prevent the Language Model from being Overconfident in Neural Machine Translation

[Mengqi Miao](#) | [Fandong Meng](#) | [Yijin Liu](#) | [Xiao-Hua Zhou](#) | [Jie Zhou](#)

[pdfbibabs](#)

Towards Emotional Support Dialog Systems

[Siyang Liu](#) | [Chujie Zheng](#) | [Orianna Demasi](#) | [Sahand Sabour](#) | [Yu Li](#) | [Zhou Yu](#) | [Yong Jiang](#) | [Minlie Huang](#)

[pdfbibabs](#)

Novel Slot Detection: A Benchmark for Discovering Unknown Slot Types in the Task-Oriented Dialogue System

[Yanan Wu](#) | [Zhiyuan Zeng](#) | [Keping He](#) | [Hong Xu](#) | [Yuanmeng Yan](#) | [Huixing Jiang](#) | [Weiran Xu](#)

[pdfbibabs](#)

GTM: A Generative Triple-wise Model for Conversational Question Generation

[Lei Shen](#) | [Fandong Meng](#) | [Jinchao Zhang](#) | [Yang Feng](#) | [Jie Zhou](#)

[pdfbibabs](#)

Diversifying Dialog Generation via Adaptive Label Smoothing

[Yida Wang](#) | [Yinhe Zheng](#) | [Yong Jiang](#) | [Minlie Huang](#)

[pdfbibabs](#)

Out-of-Scope Intent Detection with Self-Supervision and Discriminative Training

[Li-Ming Zhan](#) | [Haowen Liang](#) | [Bo Liu](#) | [Lu Fan](#) | [Xiao-Ming Wu](#) | [Albert Y.S. Lam](#)

[pdfbibabs](#)

Document-level Event Extraction via Heterogeneous Graph-based Interaction Model with a Tracker

[Runxin Xu](#) | [Tianyu Liu](#) | [Lei Li](#) | [Baobao Chang](#)

[pdfbibabs](#)

Nested Named Entity Recognition via Explicitly Excluding the Influence of the Best Path

[Yiran Wang](#) | [Hiroyuki Shindo](#) | [Yuji Matsumoto](#) | [Taro Watanabe](#)

[pdfbibabs](#)

LearnDA: Learnable Knowledge-Guided Data Augmentation for Event Causality Identification

[Xinyu Zuo](#) | [Pengfei Cao](#) | [Yubo Chen](#) | [Kang Liu](#) | [Jun Zhao](#) | [Weihua Peng](#) | [Yuguang Chen](#)

[pdfbibabs](#)

Revisiting the Negative Data of Distantly Supervised Relation Extraction

[Chenhao Xie](#) | [Jiaqing Liang](#) | [Jingping Liu](#) | [Chengsong Huang](#) | [Wenhao Huang](#) | [Yanghua Xiao](#)

[pdfbibabs](#)

Knowing the No-match: Entity Alignment with Dangling Cases

[Zequn Sun](#) | [Muhao Chen](#) | [Wei Hu](#)

[pdfbibabs](#)

Superbizarre Is Not Superb: Derivational Morphology Improves BERT's Interpretation of Complex Words

[Valentin Hofmann](#) | [Janet Pierrehumbert](#) | [Hinrich Schütze](#)

[pdfbibabs](#)

BERT is to NLP what AlexNet is to CV: Can Pre-Trained Language Models Identify Analogies?

[Asahi Ushio](#) | [Luis Espinosa Anke](#) | [Steven Schockaert](#) | [Jose Camacho-Collados](#)

[pdfbibabs](#)

Exploring the Representation of Word Meanings in Context: A Case Study on Homonymy and Synonymy

[Marcos Garcia](#)

[pdfbibabs](#)

Measuring Fine-Grained Domain Relevance of Terms: A Hierarchical Core-Fringe Approach

[Jie Huang](#) | [Kevin Chang](#) | [JinJun Xiong](#) | [Wen-mei Hwu](#)

[pdfbibabs](#)

HERALD: An Annotation Efficient Method to Detect User Disengagement in Social Conversations

[Weixin Liang](#) | [Kai-Hui Liang](#) | [Zhou Yu](#)

[pdfbibabs](#)

Value-Agnostic Conversational Semantic Parsing

[Emmanouil Antonios Platanios](#) | [Adam Pauls](#) | [Subhro Roy](#) | [Yuchen Zhang](#) | [Alexander Kyte](#) | [Alan Guo](#) | [Sam Thomson](#) | [Jayant Krishnamurthy](#) | [Jason Wolfe](#) | [Jacob Andreas](#) | [Dan Klein](#)

[pdfbibabs](#)

MPC-BERT: A Pre-Trained Language Model for Multi-Party Conversation Understanding

[Jia-Chen Gu](#) | [Chongyang Tao](#) | [Zenhua Ling](#) | [Can Xu](#) | [Xiubo Geng](#) | [Daxin Jiang](#)

[pdfbibabs](#)

Best of Both Worlds: Making High Accuracy Non-incremental Transformer-based Disfluency Detection Incremental

[Morteza Rohanian](#) | [Julian Hough](#)

[pdfbibabs](#)

NeuralWOZ: Learning to Collect Task-Oriented Dialogue via Model-Based Simulation

[Sungdong Kim](#) | [Minsuk Chang](#) | [Sang-Woo Lee](#)

[pdfbibabs](#)

CDRNN: Discovering Complex Dynamics in Human Language Processing

[Cory Shain](#)

[pdfbibabs](#)

Structural Guidance for Transformer Language Models

[Peng Qian](#) | [Tahira Naseem](#) | [Roger Levy](#) | [Ramón Fernandez Astudillo](#)

[pdfbibabs](#)

Surprisal Estimators for Human Reading Times Need Character Models

[Byung-Doh Oh](#) | [Christian Clark](#) | [William Schuler](#)

[pdfbibabs](#)

CogAlign: Learning to Align Textual Neural Representations to Cognitive Language Processing Signals

[Yuqi Ren](#) | [Deyi Xiong](#)

[pdfbibabs](#)

Self-Attention Networks Can Process Bounded Hierarchical Languages

[Shunyu Yao](#) | [Binghui Peng](#) | [Christos Papadimitriou](#) | [Karthik Narasimhan](#)

[pdfbibabs](#)

TextSETTR: Few-Shot Text Style Extraction and Tunable Targeted Restyling

[Parker Riley](#) | [Noah Constant](#) | [Mandy Guo](#) | [Girish Kumar](#) | [David Uthus](#) | [Zarana Parekh](#)

[pdfbibabs](#)

H-Transformer-1D: Fast One-Dimensional Hierarchical Attention for Sequences

[Zhenhai Zhu](#) | [Radu Soricut](#)

[pdfbibabs](#)

Making Pre-trained Language Models Better Few-shot Learners

[Tianyu Gao](#) | [Adam Fisch](#) | [Danqi Chen](#)

[pdfbibabs](#)

A Sweet Rabbit Hole by DARCY: Using Honey Pots to Detect Universal Trigger's Adversarial Attacks

[Thai Le](#) | [Noseong Park](#) | [Dongwon Lee](#)

[pdfbibabs](#)

Towards Propagation Uncertainty: Edge-enhanced Bayesian Graph Convolutional Networks for Rumor Detection

[Lingwei Wei](#) | [Dou Hu](#) | [Wei Zhou](#) | [Zhaojuan Yue](#) | [Songlin Hu](#)

[pdfbibabs](#)

Label-Specific Dual Graph Neural Network for Multi-Label Text Classification

[Qianwen Ma](#) | [Chunyuan Yuan](#) | [Wei Zhou](#) | [Songlin Hu](#)

[pdfbibabs](#)

TAN-NTM: Topic Attention Networks for Neural Topic Modeling

[Madhur Panwar](#) | [Shashank Shailabh](#) | [Milan Aggarwal](#) | [Balaji Krishnamurthy](#)

[pdfbibabs](#)

Cross-language Sentence Selection via Data Augmentation and Rationale Training

[Yanda Chen](#) | [Chris Kedzie](#) | [Suraj Nair](#) | [Petra Galuscakova](#) | [Rui Zhang](#) | [Douglas Oard](#) | [Kathleen McKeown](#)

[pdfbibabs](#)

A Neural Model for Joint Document and Snippet Ranking in Question Answering for Large Document Collections

[Dimitris Pappas](#) | [Ion Androutsopoulos](#)

[pdfbibabs](#)

W-RST: Towards a Weighted RST-style Discourse Framework

[Patrick Huber](#) | [Wen Xiao](#) | [Giuseppe Carenini](#)

[pdfbibabs](#)

ABCD: A Graph Framework to Convert Complex Sentences to a Covering Set of Simple Sentences

[Yanjun Gao](#) | [Ting-Hao Huang](#) | [Rebecca J. Passonneau](#)

[pdfbibabs](#)

Which Linguist Invented the Lightbulb? Presupposition Verification for Question-Answering

[Najoung Kim](#) | [Ellie Pavlick](#) | [Burcu Karagol Ayan](#) | [Deepak Ramachandran](#)

[pdfbibabs](#)

Adversarial Learning for Discourse Rhetorical Structure Parsing

[Longyin Zhang](#) | [Fang Kong](#) | [Guodong Zhou](#)

[pdfbibabs](#)

Exploring Discourse Structures for Argument Impact Classification

[Xin Liu](#) | [Jiefu Ou](#) | [Yangqiu Song](#) | [Xin Jiang](#)

[pdfbibabs](#)

Point, Disambiguate and Copy: Incorporating Bilingual Dictionaries for Neural Machine Translation

[Tong Zhang](#) | [Long Zhang](#) | [Wei Ye](#) | [Bo Li](#) | [Jinan Sun](#) | [Xiaoyu Zhu](#) | [Wen Zhao](#) | [Shikun Zhang](#)

[pdfbibabs](#)

VECO: Variable and Flexible Cross-lingual Pre-training for Language Understanding and Generation

[Fuli Luo](#) | [Wei Wang](#) | [Jiahao Liu](#) | [Yijia Liu](#) | [Bin Bi](#) | [Songfang Huang](#) | [Fei Huang](#) | [Luo Si](#)

[pdfbibabs](#)

A unified approach to sentence segmentation of punctuated text in many languages

[Rachel Wicks](#) | [Matt Post](#)

[pdfbibabs](#)

Towards User-Driven Neural Machine Translation

[Huan Lin](#) | [Liang Yao](#) | [Baosong Yang](#) | [Dayiheng Liu](#) | [Haibo Zhang](#) | [Weihua Luo](#) | [Degen Huang](#) | [Jinsong Su](#)

[pdfbibabs](#)

[End-to-End Lexically Constrained Machine Translation for Morphologically Rich Languages](#)

[Josef Jon](#) | [João Paulo Aires](#) | [Dusan Varis](#) | [Ondřej Bojar](#)

[pdfbibabs](#)

[Handling Extreme Class Imbalance in Technical Logbook Datasets](#)

[Farhad Akhbardeh](#) | [Cecilia Ovesdotter Alm](#) | [Marcos Zampieri](#) | [Travis Desell](#)

[pdfbibabs](#)

[ILDC for CJPE: Indian Legal Documents Corpus for Court Judgment Prediction and Explanation](#)

[Vijit Malik](#) | [Rishabh Sanjay](#) | [Shubham Kumar Nigam](#) | [Kripabandhu Ghosh](#) | [Shouvik Kumar Guha](#) | [Arnab Bhattacharya](#) | [Ashutosh Modi](#)

[pdfbibabs](#)

[Supporting Cognitive and Emotional Empathic Writing of Students](#)

[Thiemo Wambganss](#) | [Christina Niklaus](#) | [Matthias Söllner](#) | [Siegfried Handschuh](#) | [Jan Marco Leimeister](#)

[pdfbibabs](#)

[Dual Reader-Parser on Hybrid Textual and Tabular Evidence for Open Domain Question Answering](#)

[Alexander Hanbo Li](#) | [Patrick Ng](#) | [Peng Xu](#) | [Henghui Zhu](#) | [Zhiguo Wang](#) | [Bing Xiang](#)

[pdfbibabs](#)

[Generation-Augmented Retrieval for Open-Domain Question Answering](#)

[Yuning Mao](#) | [Pengcheng He](#) | [Xiaodong Liu](#) | [Yelong Shen](#) | [Jianfeng Gao](#) | [Jiawei Han](#) | [Weizhu Chen](#)

[pdfbibabs](#)

[Check It Again: Progressive Visual Question Answering via Visual Entailment](#)

[Qingyi Si](#) | [Zheng Lin](#) | [Mingyu Zheng](#) | [Peng Fu](#) | [Weiping Wang](#)

[pdfbibabs](#)

[A Mutual Information Maximization Approach for the Spurious Solution Problem in Weakly Supervised Question Answering](#)

[Zhihong Shao](#) | [Lifeng Shang](#) | [Qun Liu](#) | [Minlie Huang](#)

[pdfbibabs](#)

[Breaking Down Walls of Text: How Can NLP Benefit Consumer Privacy?](#)

[Abhilasha Ravichander](#) | [Alan W Black](#) | [Thomas Norton](#) | [Shomir Wilson](#) | [Norman Sadeh](#)

[pdfbibabs](#)

[Supporting Land Reuse of Former Open Pit Mining Sites using Text Classification and Active Learning](#)

[Christopher Schröder](#) | [Kim Bürgl](#) | [Yves Annanias](#) | [Andreas Niekler](#) | [Lydia Müller](#) | [Daniel Wiegrefe](#) | [Christian Bender](#) | [Christoph Mengs](#) | [Gerik Scheuermann](#) | [Gerhard Heyer](#)

[pdfbibabs](#)

[Reliability Testing for Natural Language Processing Systems](#)

[Samson Tan](#) | [Shafiq Joty](#) | [Kathy Baxter](#) | [Araz Taeihagh](#) | [Gregory A. Bennett](#) | [Min-Yen Kan](#)

[pdfbibabs](#)

[Learning Language and Multimodal Privacy-Preserving Markers of Mood from Mobile Data](#)

[Paul Pu Liang](#) | [Terrance Liu](#) | [Anna Cai](#) | [Michał Muszynski](#) | [Ryo Ishii](#) | [Nick Allen](#) | [Randy Auerbach](#) | [David Brent](#) | [Ruslan Salakhutdinov](#) | [Louis-Philippe Morency](#)

[pdfbibabs](#)

Anonymisation Models for Text Data: State of the art, Challenges and Future Directions

[Pierre Lison](#) | [Ildikó Pilán](#) | [David Sanchez](#) | [Montserrat Batet](#) | [Lilja Øvrelid](#)

[pdfbibabs](#)

End-to-End AMR Coreference Resolution

[Qiankun Fu](#) | [Linfeng Song](#) | [Wenyu Du](#) | [Yue Zhang](#)

[pdfbibabs](#)

How is BERT surprised? Layerwise detection of linguistic anomalies

[Bai Li](#) | [Zining Zhu](#) | [Guillaume Thomas](#) | [Yang Xu](#) | [Frank Rudzicz](#)

[pdfbibabs](#)

Psycholinguistic Tripartite Graph Network for Personality Detection

[Tao Yang](#) | [Feifan Yang](#) | [Haolan Ouyang](#) | [Xiaojun Quan](#)

[pdfbibabs](#)

Verb Metaphor Detection via Contextual Relation Learning

[Wei Song](#) | [Shuhui Zhou](#) | [Ruiji Fu](#) | [Ting Liu](#) | [Lizhen Liu](#)

[pdfbibabs](#)

Improving Speech Translation by Understanding and Learning from the Auxiliary Text Translation Task

[Yun Tang](#) | [Juan Pino](#) | [Xian Li](#) | [Changhan Wang](#) | [Dmitriy Genzel](#)

[pdfbibabs](#)

Probing Toxic Content in Large Pre-Trained Language Models

[Nedjma Ousidhoum](#) | [Xinran Zhao](#) | [Tianqing Fang](#) | [Yangqiu Song](#) | [Dit-Yan Yeung](#)

[pdfbibabs](#)

Societal Biases in Language Generation: Progress and Challenges

[Emily Sheng](#) | [Kai-Wei Chang](#) | [Prem Natarajan](#) | [Nanyun Peng](#)

[pdfbibabs](#)

Reservoir Transformers

[Sheng Shen](#) | [Alexei Baeviski](#) | [Ari Morcos](#) | [Kurt Keutzer](#) | [Michael Auli](#) | [Douwe Kiela](#)

[pdfbibabs](#)

Subsequence Based Deep Active Learning for Named Entity Recognition

[Puria Radmard](#) | [Yassir Fathullah](#) | [Aldo Lipani](#)

[pdfbibabs](#)

Convolutions and Self-Attention: Re-interpreting Relative Positions in Pre-trained Language Models

[Tyler Chang](#) | [Yifan Xu](#) | [Weijian Xu](#) | [Zhuowen Tu](#)

[pdfbibabs](#)

BinaryBERT: Pushing the Limit of BERT Quantization

[Haoli Bai](#) | [Wei Zhang](#) | [Lu Hou](#) | [Lifeng Shang](#) | [Jin Jin](#) | [Xin Jiang](#) | [Qun Liu](#) | [Michael Lyu](#) | [Irwin King](#)

[pdfbibabs](#)

Are Pretrained Convolutions Better than Pretrained Transformers?

[Yi Tay](#) | [Mostafa Dehghani](#) | [Jai Prakash Gupta](#) | [Vamsi Aribandi](#) | [Dara Bahri](#) | [Zhen Qin](#) | [Donald Metzler](#)

[pdfbibabs](#)

[PairRE: Knowledge Graph Embeddings via Paired Relation Vectors](#)

[Linlin Chao](#) | [Jianshan He](#) | [Taifeng Wang](#) | [Wei Chu](#)

[pdfbibabs](#)

[Hierarchy-aware Label Semantics Matching Network for Hierarchical Text Classification](#)

[Haibin Chen](#) | [Qianli Ma](#) | [Zhenxi Lin](#) | [Jiangyue Yan](#)

[pdfbibabs](#)

[HiddenCut: Simple Data Augmentation for Natural Language Understanding with Better Generalizability](#)

[Jiaao Chen](#) | [Dinghan Shen](#) | [Weizhu Chen](#) | [Diyi Yang](#)

[pdfbibabs](#)

[Neural Stylistic Response Generation with Disentangled Latent Variables](#)

[Qingfu Zhu](#) | [Wei-Nan Zhang](#) | [Ting Liu](#) | [William Yang Wang](#)

[pdfbibabs](#)

[Intent Classification and Slot Filling for Privacy Policies](#)

[Wasi Ahmad](#) | [Jianfeng Chi](#) | [Tu Le](#) | [Thomas Norton](#) | [Yuan Tian](#) | [Kai-Wei Chang](#)

[pdfbibabs](#)

[RADDLE: An Evaluation Benchmark and Analysis Platform for Robust Task-oriented Dialog Systems](#)

[Baolin Peng](#) | [Chunyu Li](#) | [Zhu Zhang](#) | [Chenguang Zhu](#) | [Jinchao Li](#) | [Jianfeng Gao](#)

[pdfbibabs](#)

[Semantic Representation for Dialogue Modeling](#)

[Xuefeng Bai](#) | [Yulong Chen](#) | [Linfeng Song](#) | [Yue Zhang](#)

[pdfbibabs](#)

[A Pre-training Strategy for Zero-Resource Response Selection in Knowledge-Grounded Conversations](#)

[Chongyang Tao](#) | [Changyu Chen](#) | [Jiazhan Feng](#) | [Ji-Rong Wen](#) | [Rui Yan](#)

[pdfbibabs](#)

[Dependency-driven Relation Extraction with Attentive Graph Convolutional Networks](#)

[Yuanhe Tian](#) | [Guimin Chen](#) | [Yan Song](#) | [Xiang Wan](#)

[pdfbibabs](#)

[Evaluating Entity Disambiguation and the Role of Popularity in Retrieval-Based NLP](#)

[Anthony Chen](#) | [Pallavi Gudipati](#) | [Shayne Longpre](#) | [Xiao Ling](#) | [Sameer Singh](#)

[pdfbibabs](#)

[Evaluation Examples are not Equally Informative: How should that change NLP Leaderboards?](#)

[Pedro Rodriguez](#) | [Joe Barrow](#) | [Alexander Miserlis Hoyle](#) | [John P. Lalor](#) | [Robin Jia](#) | [Jordan Boyd-Graber](#)

[pdfbibabs](#)

[Claim Matching Beyond English to Scale Global Fact-Checking](#)

[Ashkan Kazemi](#) | [Kiran Garimella](#) | [Devin Gaffney](#) | [Scott Hale](#)

[pdfbibabs](#)

[SemFace: Pre-training Encoder and Decoder with a Semantic Interface for Neural Machine Translation](#)

[Shuo Ren](#) | [Long Zhou](#) | [Shujie Liu](#) | [Furu Wei](#) | [Ming Zhou](#) | [Shuai Ma](#)

[pdfbibabs](#)

Energy-Based Reranking: Improving Neural Machine Translation Using Energy-Based Models

[Sumanta Bhattacharyya](#) | [Amirmohammad Rooshenas](#) | [Subhajit Naskar](#) | [Simeng Sun](#) | [Mohit Iyer](#) | [Andrew McCallum](#)

[pdfbibabs](#)

Syntax-augmented Multilingual BERT for Cross-lingual Transfer

[Wasi Ahmad](#) | [Haoran Li](#) | [Kai-Wei Chang](#) | [Yashar Mehdad](#)

[pdfbibabs](#)

How to Adapt Your Pretrained Multilingual Model to 1600 Languages

[Abteen Ebrahimi](#) | [Katharina Kann](#)

[pdfbibabs](#)

Weakly Supervised Named Entity Tagging with Learnable Logical Rules

[Jiacheng Li](#) | [Haibo Ding](#) | [Jingbo Shang](#) | [Julian McAuley](#) | [Zhe Feng](#)

[pdfbibabs](#)

Prefix-Tuning: Optimizing Continuous Prompts for Generation

[Xiang Lisa Li](#) | [Percy Liang](#)

[pdfbibabs](#)

One2Set: Generating Diverse Keyphrases as a Set

[Jiacheng Ye](#) | [Tao Gui](#) | [Yichao Luo](#) | [Yige Xu](#) | [Qi Zhang](#)

[pdfbibabs](#)

Continuous Language Generative Flow

[Zineng Tang](#) | [Shiyue Zhang](#) | [Hyounghun Kim](#) | [Mohit Bansal](#)

[pdfbibabs](#)

TWAG: A Topic-Guided Wikipedia Abstract Generator

[Fangwei Zhu](#) | [Shangqing Tu](#) | [Jiaxin Shi](#) | [Juanzi Li](#) | [Lei Hou](#) | [Tong Cui](#)

[pdfbibabs](#)

ForecastQA: A Question Answering Challenge for Event Forecasting with Temporal Text Data

[Woojeong Jin](#) | [Rahul Khanna](#) | [Suji Kim](#) | [Dong-Ho Lee](#) | [Fred Morstatter](#) | [Aram Galstyan](#) | [Xiang Ren](#)

[pdfbibabs](#)

Recursive Tree-Structured Self-Attention for Answer Sentence Selection

[Khalil Mrini](#) | [Emilia Farcas](#) | [Ndapa Nakashole](#)

[pdfbibabs](#)

How Knowledge Graph and Attention Help? A Qualitative Analysis into Bag-level Relation Extraction

[Zikun Hu](#) | [Yixin Cao](#) | [Lifu Huang](#) | [Tat-Seng Chua](#)

[pdfbibabs](#)

Trigger is Not Sufficient: Exploiting Frame-aware Knowledge for Implicit Event Argument Extraction

[Kaiwen Wei](#) | [Xian Sun](#) | [Zequn Zhang](#) | [Jingyuan Zhang](#) | [Guo Zhi](#) | [Li Jin](#)

[pdfbibabs](#)

Element Intervention for Open Relation Extraction

[Fangchao Liu](#) | [Lingyong Yan](#) | [Hongyu Lin](#) | [Xianpei Han](#) | [Le Sun](#)

[pdfbibabs](#)

[AdaTag: Multi-Attribute Value Extraction from Product Profiles with Adaptive Decoding](#)

[Jun Yan](#) | [Nasser Zalmout](#) | [Yan Liang](#) | [Christan Grant](#) | [Xiang Ren](#) | [Xin Luna Dong](#)

[pdfbibabs](#)

[CoRI: Collective Relation Integration with Data Augmentation for Open Information Extraction](#)

[Zhengbao Jiang](#) | [Jialong Han](#) | [Bunyamin Sisman](#) | [Xin Luna Dong](#)

[pdfbibabs](#)

[Benchmarking Scalable Methods for Streaming Cross Document Entity Coreference](#)

[Robert L Logan IV](#) | [Andrew McCallum](#) | [Sameer Singh](#) | [Dan Bikel](#)

[pdfbibabs](#)

[Search from History and Reason for Future: Two-stage Reasoning on Temporal Knowledge Graphs](#)

[Zixuan Li](#) | [Xiaolong Jin](#) | [Saiping Guan](#) | [Wei Li](#) | [Jiafeng Guo](#) | [Yuanzhuo Wang](#) | [Xueqi Cheng](#)

[pdfbibabs](#)

[Employing Argumentation Knowledge Graphs for Neural Argument Generation](#)

[Khalid Al Khatib](#) | [Lukas Trautner](#) | [Henning Wachsmuth](#) | [Yufang Hou](#) | [Benno Stein](#)

[pdfbibabs](#)

[Learning Span-Level Interactions for Aspect Sentiment Triplet Extraction](#)

[Lu Xu](#) | [Yew Ken Chia](#) | [Lidong Bing](#)

[pdfbibabs](#)

[On Compositional Generalization of Neural Machine Translation](#)

[Yafu Li](#) | [Yongjing Yin](#) | [Yulong Chen](#) | [Yue Zhang](#)

[pdfbibabs](#)

[Mask-Align: Self-Supervised Neural Word Alignment](#)

[Chi Chen](#) | [Maosong Sun](#) | [Yang Liu](#)

[pdfbibabs](#)

[GWLAN: General Word-Level Autocompletion for Computer-Aided Translation](#)

[Huayang Li](#) | [Lemao Liu](#) | [Guoping Huang](#) | [Shuming Shi](#)

[pdfbibabs](#)

[De-biasing Distantly Supervised Named Entity Recognition via Causal Intervention](#)

[Wenkai Zhang](#) | [Hongyu Lin](#) | [Xianpei Han](#) | [Le Sun](#)

[pdfbibabs](#)

[A Span-Based Model for Joint Overlapped and Discontinuous Named Entity Recognition](#)

[Fei Li](#) | [ZhiChao Lin](#) | [Meishan Zhang](#) | [Donghong Ji](#)

[pdfbibabs](#)

[MLBiNet: A Cross-Sentence Collective Event Detection Network](#)

[Dongfang Lou](#) | [Zhilin Liao](#) | [Shumin Deng](#) | [Ningyu Zhang](#) | [Huajun Chen](#)

[pdfbibabs](#)

[Exploiting Document Structures and Cluster Consistencies for Event Coreference Resolution](#)

[Hieu Minh Tran](#) | [Duy Phung](#) | [Thien Huu Nguyen](#)

[pdfbibabs](#)

[StereoRel: Relational Triple Extraction from a Stereoscopic Perspective](#)

[Xuetao Tian](#) | [Liping Jing](#) | [Lu He](#) | [Feng Liu](#)

[pdfbibabs](#)

Knowledge-Enriched Event Causality Identification via Latent Structure Induction Networks

[Pengfei Cao](#) | [Xinyu Zuo](#) | [Yubo Chen](#) | [Kang Liu](#) | [Jun Zhao](#) | [Yuguang Chen](#) | [Weihua Peng](#)

[pdfbibabs](#)

Turn the Combination Lock: Learnable Textual Backdoor Attacks via Word Substitution

[Fanchao Qi](#) | [Yuan Yao](#) | [Sophia Xu](#) | [Zhiyuan Liu](#) | [Maosong Sun](#)

[pdfbibabs](#)

Parameter-Efficient Transfer Learning with Diff Pruning

[Demi Guo](#) | [Alexander Rush](#) | [Yoon Kim](#)

[pdfbibabs](#)

R2D2: Recursive Transformer based on Differentiable Tree for Interpretable Hierarchical Language Modeling

[Xiang Hu](#) | [Haitao Mi](#) | [Zujie Wen](#) | [Yafang Wang](#) | [Yi Su](#) | [Jing Zheng](#) | [Gerard de Melo](#)

[pdfbibabs](#)

Risk Minimization for Zero-shot Sequence Labeling

[Zechuan Hu](#) | [Yong Jiang](#) | [Nguyen Bach](#) | [Tao Wang](#) | [Zhongqiang Huang](#) | [Fei Huang](#) | [Kewei Tu](#)

[pdfbibabs](#)

WARP: Word-level Adversarial ReProgramming

[Karen Hambardzumyan](#) | [Hrant Khachatrian](#) | [Jonathan May](#)

[pdfbibabs](#)

Lexicon Learning for Few Shot Sequence Modeling

[Ekin Akyurek](#) | [Jacob Andreas](#)

[pdfbibabs](#)

Personalized Transformer for Explainable Recommendation

[Lei Li](#) | [Yongfeng Zhang](#) | [Li Chen](#)

[pdfbibabs](#)

Generating SOAP Notes from Doctor-Patient Conversations Using Modular Summarization Techniques

[Kundan Krishna](#) | [Sopan Khosla](#) | [Jeffrey Bigham](#) | [Zachary C. Lipton](#)

[pdfbibabs](#)

Tail-to-Tail Non-Autoregressive Sequence Prediction for Chinese Grammatical Error Correction

[Piji Li](#) | [Shuming Shi](#)

[pdfbibabs](#)

Early Detection of Sexual Predators in Chats

[Matthias Vogt](#) | [Ulf Leser](#) | [Alan Akbik](#)

[pdfbibabs](#)

Writing by Memorizing: Hierarchical Retrieval-based Medical Report Generation

[Xingyi Yang](#) | [Muchao Ye](#) | [Quanzeng You](#) | [Fenglong Ma](#)

[pdfbibabs](#)

Concept-Based Label Embedding via Dynamic Routing for Hierarchical Text Classification

[Xuepeng Wang](#) | [Li Zhao](#) | [Bing Liu](#) | [Tao Chen](#) | [Feng Zhang](#) | [Di Wang](#)

[pdfbibabs](#)

VisualSparta: An Embarrassingly Simple Approach to Large-scale Text-to-Image Search with Weighted Bag-of-words

[Xiaopeng Lu](#) | [Tiancheng Zhao](#) | [Kyunsoong Lee](#)

[pdfbibabs](#)

Few-Shot Text Ranking with Meta Adapted Synthetic Weak Supervision

[Si Sun](#) | [Yingzhuo Qian](#) | [Zhenghao Liu](#) | [Chenyan Xiong](#) | [Kaitao Zhang](#) | [Jie Bao](#) | [Zhiyuan Liu](#) | [Paul Bennett](#)

[pdfbibabs](#)

Semi-Supervised Text Classification with Balanced Deep Representation Distributions

[Changchun Li](#) | [Ximing Li](#) | [Jihong Ouyang](#)

[pdfbibabs](#)

Improving Document Representations by Generating Pseudo Query Embeddings for Dense Retrieval

[Hongyin Tang](#) | [Xingwu Sun](#) | [Beihong Jin](#) | [Jingang Wang](#) | [Fuzheng Zhang](#) | [Wei Wu](#)

[pdfbibabs](#)

ConSERT: A Contrastive Framework for Self-Supervised Sentence Representation Transfer

[Yuanmeng Yan](#) | [Rumei Li](#) | [Sirui Wang](#) | [Fuzheng Zhang](#) | [Wei Wu](#) | [Weiran Xu](#)

[pdfbibabs](#)

Exploring Dynamic Selection of Branch Expansion Orders for Code Generation

[Hui Jiang](#) | [Chulun Zhou](#) | [Fandong Meng](#) | [Biao Zhang](#) | [Jie Zhou](#) | [Degen Huang](#) | [Qingqiang Wu](#) | [Jinsong Su](#)

[pdfbibabs](#)

COINS: Dynamically Generating Contextualized Inference Rules for Narrative Story Completion

[Debjit Paul](#) | [Anette Frank](#)

[pdfbibabs](#)

Reasoning over Entity-Action-Location Graph for Procedural Text Understanding

[Hao Huang](#) | [Xiubo Geng](#) | [Jian Pei](#) | [Guodong Long](#) | [Daxin Jiang](#)

[pdfbibabs](#)

From Paraphrasing to Semantic Parsing: Unsupervised Semantic Parsing via Synchronous Semantic Decoding

[Shan Wu](#) | [Bo Chen](#) | [Chunlei Xin](#) | [Xianpei Han](#) | [Le Sun](#) | [Weipeng Zhang](#) | [Jiansong Chen](#) | [Fan Yang](#) | [Xunliang Cai](#)

[pdfbibabs](#)

Pre-training Universal Language Representation

[Yian Li](#) | [Hai Zhao](#)

[pdfbibabs](#)

Structural Pre-training for Dialogue Comprehension

[Zhuosheng Zhang](#) | [Hai Zhao](#)

[pdfbibabs](#)

AutoTinyBERT: Automatic Hyper-parameter Optimization for Efficient Pre-trained Language Models

[Yichun Yin](#) | [Cheng Chen](#) | [Lifeng Shang](#) | [Xin Jiang](#) | [Xiao Chen](#) | [Qun Liu](#)

[pdfbibabs](#)

[Data Augmentation with Adversarial Training for Cross-Lingual NLI](#)

[Xin Dong](#) | [Yaxin Zhu](#) | [Zuohui Fu](#) | [Dongkuan Xu](#) | [Gerard de Melo](#)

[pdfbibabs](#)

[Bootstrapped Unsupervised Sentence Representation Learning](#)

[Yan Zhang](#) | [Ruidan He](#) | [Zuozhu Liu](#) | [Lidong Bing](#) | [Haizhou Li](#)

[pdfbibabs](#)

[Learning Event Graph Knowledge for Abductive Reasoning](#)

[Li Du](#) | [Xiao Ding](#) | [Ting Liu](#) | [Bing Qin](#)

[pdfbibabs](#)

[A Cognitive Regularizer for Language Modeling](#)

[Jason Wei](#) | [Clara Meister](#) | [Ryan Cotterell](#)

[pdfbibabs](#)

[Lower Perplexity is Not Always Human-Like](#)

[Tatsuki Kuribayashi](#) | [Yohei Oseki](#) | [Takumi Ito](#) | [Ryo Yoshida](#) | [Masayuki Asahara](#) | [Kentarō Inui](#)

[pdfbibabs](#)

[Word Sense Disambiguation: Towards Interactive Context Exploitation from Both Word and Sense Perspectives](#)

[Ming Wang](#) | [Yinglin Wang](#)

[pdfbibabs](#)

[A Knowledge-Guided Framework for Frame Identification](#)

[Xuefeng Su](#) | [Ru Li](#) | [Xiaoli Li](#) | [Jeff Z. Pan](#) | [Hu Zhang](#) | [Qinghua Chai](#) | [Xiaoqi Han](#)

[pdfbibabs](#)

[Obtaining Better Static Word Embeddings Using Contextual Embedding Models](#)

[Prakhar Gupta](#) | [Martin Jaggi](#)

[pdfbibabs](#)

[Meta-Learning with Variational Semantic Memory for Word Sense Disambiguation](#)

[Yingjun Du](#) | [Nithin Holla](#) | [Xiantong Zhen](#) | [Cees Snoek](#) | [Ekaterina Shutova](#)

[pdfbibabs](#)

[LexFit: Lexical Fine-Tuning of Pretrained Language Models](#)

[Ivan Vulić](#) | [Edoardo Maria Ponti](#) | [Anna Korhonen](#) | [Goran Glavaš](#)

[pdfbibabs](#)

[Text-Free Image-to-Speech Synthesis Using Learned Segmental Units](#)

[Wei-Ning Hsu](#) | [David Harwath](#) | [Tyler Miller](#) | [Christopher Song](#) | [James Glass](#)

[pdfbibabs](#)

[CTFN: Hierarchical Learning for Multimodal Sentiment Analysis Using Coupled-Translation Fusion Network](#)

[Jiajia Tang](#) | [Kang Li](#) | [Xuanyu Jin](#) | [Andrzej Cichocki](#) | [Qibin Zhao](#) | [Wanzeng Kong](#)

[pdfbibabs](#)

[Positional Artefacts Propagate Through Masked Language Model Embeddings](#)

[Ziyang Luo](#) | [Artur Kulmizev](#) | [Xiaoxi Mao](#)

[pdfbibabs](#)

[Language Model Evaluation Beyond Perplexity](#)

[Clara Meister](#) | [Ryan Cotterell](#)

[pdfbibabs](#)

[Learning to Explain: Generating Stable Explanations Fast](#)

[Xuelin Situ](#) | [Ingrid Zukerman](#) | [Cecile Paris](#) | [Sameen Maruf](#) | [Gholamreza Haffari](#)

[pdfbibabs](#)

[StereoSet: Measuring stereotypical bias in pretrained language models](#)

[Moin Nadeem](#) | [Anna Bethke](#) | [Siva Reddy](#)

[pdfbibabs](#)

[Alignment Rationale for Natural Language Inference](#)

[Zhongtao Jiang](#) | [Yuanzhe Zhang](#) | [Zhao Yang](#) | [Jun Zhao](#) | [Kang Liu](#)

[pdfbibabs](#)

[Enabling Lightweight Fine-tuning for Pre-trained Language Model Compression based on Matrix Product Operators](#)

[Peiyu Liu](#) | [Ze-Feng Gao](#) | [Wayne Xin Zhao](#) | [Zhi-Yuan Xie](#) | [Zhong-Yi Lu](#) | [Ji-Rong Wen](#)

[pdfbibabs](#)

[On Sample Based Explanation Methods for NLP: Faithfulness, Efficiency and Semantic Evaluation](#)

[Wei Zhang](#) | [Ziming Huang](#) | [Yada Zhu](#) | [Guangnan Ye](#) | [Xiaodong Cui](#) | [Fan Zhang](#)

[pdfbibabs](#)

[Syntax-Enhanced Pre-trained Model](#)

[Zenan Xu](#) | [Daya Guo](#) | [Duyu Tang](#) | [Qinliang Su](#) | [Linjun Shou](#) | [Ming Gong](#) | [Wanjuan Zhong](#) | [Xiaojuan Quan](#) | [Daxin Jiang](#) | [Nan Duan](#)

[pdfbibabs](#)

[Matching Distributions between Model and Data: Cross-domain Knowledge Distillation for Unsupervised Domain Adaptation](#)

[Bo Zhang](#) | [Xiaoming Zhang](#) | [Yun Liu](#) | [Lei Cheng](#) | [Zhoujun Li](#)

[pdfbibabs](#)

[Counterfactual Inference for Text Classification Debiasing](#)

[Chen Qian](#) | [Fuli Feng](#) | [Lijie Wen](#) | [Chunping Ma](#) | [Pengjun Xie](#)

[pdfbibabs](#)

[HieRec: Hierarchical User Interest Modeling for Personalized News Recommendation](#)

[Tao Qi](#) | [Fangzhao Wu](#) | [Chuhan Wu](#) | [Peiru Yang](#) | [Yang Yu](#) | [Xing Xie](#) | [Yongfeng Huang](#)

[pdfbibabs](#)

[PP-Rec: News Recommendation with Personalized User Interest and Time-aware News Popularity](#)

[Tao Qi](#) | [Fangzhao Wu](#) | [Chuhan Wu](#) | [Yongfeng Huang](#)

[pdfbibabs](#)

[Article Reranking by Memory-Enhanced Key Sentence Matching for Detecting Previously Fact-Checked Claims](#)

[Qiang Sheng](#) | [Juan Cao](#) | [Xueyao Zhang](#) | [Xirong Li](#) | [Lei Zhong](#)

[pdfbibabs](#)

[Defense against Synonym Substitution-based Adversarial Attacks via Dirichlet Neighborhood Ensemble](#)

[Yi Zhou](#) | [Xiaoqing Zheng](#) | [Cho-Jui Hsieh](#) | [Kai-Wei Chang](#) | [Xuanjing Huang](#)

[pdfbibabs](#)

[Shortformer: Better Language Modeling using Shorter Inputs](#)

[Ofir Press](#) | [Noah A. Smith](#) | [Mike Lewis](#)

[pdfbibabs](#)

[BanditMTL: Bandit-based Multi-task Learning for Text Classification](#)

[Yuren Mao](#) | [Zekai Wang](#) | [Weiwei Liu](#) | [Xuemin Lin](#) | [Wenbin Hu](#)

[pdfbibabs](#)

[Unified Interpretation of Softmax Cross-Entropy and Negative Sampling: With Case Study for Knowledge Graph Embedding](#)

[Hidetaka Kamigaito](#) | [Katsuhiko Hayashi](#)

[pdfbibabs](#)

[De-Confounded Variational Encoder-Decoder for Logical Table-to-Text Generation](#)

[Wenqing Chen](#) | [Jidong Tian](#) | [Yitian Li](#) | [Hao He](#) | [Yaohui Jin](#)

[pdfbibabs](#)

[Rethinking Stealthiness of Backdoor Attack against NLP Models](#)

[Wenkai Yang](#) | [Yankai Lin](#) | [Peng Li](#) | [Jie Zhou](#) | [Xu Sun](#)

[pdfbibabs](#)

[Crowdsourcing Learning as Domain Adaptation: A Case Study on Named Entity Recognition](#)

[Xin Zhang](#) | [Guangwei Xu](#) | [Yueheng Sun](#) | [Meishan Zhang](#) | [Pengjun Xie](#)

[pdfbibabs](#)

[Exploring Distantly-Labeled Rationales in Neural Network Models](#)

[Quzhe Huang](#) | [Shengqi Zhu](#) | [Yansong Feng](#) | [Dongyan Zhao](#)

[pdfbibabs](#)

[Learning to Perturb Word Embeddings for Out-of-distribution QA](#)

[Seanie Lee](#) | [Minki Kang](#) | [Juho Lee](#) | [Sung Ju Hwang](#)

[pdfbibabs](#)

[Maria: A Visual Experience Powered Conversational Agent](#)

[Zujie Liang](#) | [Huang Hu](#) | [Can Xu](#) | [Chongyang Tao](#) | [Xiubo Geng](#) | [Yining Chen](#) | [Fan Liang](#) | [Daxin Jiang](#)

[pdfbibabs](#)

[A Human-machine Collaborative Framework for Evaluating Malevolence in Dialogues](#)

[Yangjun Zhang](#) | [Pengjie Ren](#) | [Maarten de Rijke](#)

[pdfbibabs](#)

[Generating Relevant and Coherent Dialogue Responses using Self-Separated Conditional Variational AutoEncoders](#)

[Bin Sun](#) | [Shaoxiong Feng](#) | [Yiwei Li](#) | [Jiamou Liu](#) | [Kan Li](#)

[pdfbibabs](#)

[Learning to Ask Conversational Questions by Optimizing Levenshtein Distance](#)

[Zhongkun Liu](#) | [Pengjie Ren](#) | [Zhumin Chen](#) | [Zhaochun Ren](#) | [Maarten de Rijke](#) | [Ming Zhou](#)

[pdfbibabs](#)

[DVD: A Diagnostic Dataset for Multi-step Reasoning in Video Grounded Dialogue](#)

[Hung Le](#) | [Chinnadhurai Sankar](#) | [Seungwhan Moon](#) | [Ahmad Beirami](#) | [Alborz Geramifard](#) | [Satwik Kottur](#)

[pdfbibabs](#)

[MMGCN: Multimodal Fusion via Deep Graph Convolution Network for Emotion Recognition in Conversation](#)

[Jingwen Hu](#) | [Yuchen Liu](#) | [Jinming Zhao](#) | [Qin Jin](#)

[pdfbibabs](#)

[DynaEval: Unifying Turn and Dialogue Level Evaluation](#)

[Chen Zhang](#) | [Yiming Chen](#) | [Luis Fernando D'Haro](#) | [Yan Zhang](#) | [Thomas Friedrichs](#) | [Grandee Lee](#) | [Haizhou Li](#)

[pdfbibabs](#)

[CoSQA: 20,000+ Web Queries for Code Search and Question Answering](#)

[Junjie Huang](#) | [Duyu Tang](#) | [Linjun Shou](#) | [Ming Gong](#) | [Ke Xu](#) | [Daxin Jiang](#) | [Ming Zhou](#) | [Nan Duan](#)

[pdfbibabs](#)

[Rewriter-Evaluator Architecture for Neural Machine Translation](#)

[Yangming Li](#) | [Kaisheng Yao](#)

[pdfbibabs](#)

[Modeling Bilingual Conversational Characteristics for Neural Chat Translation](#)

[Yunlong Liang](#) | [Fandong Meng](#) | [Yufeng Chen](#) | [Jinan Xu](#) | [Jie Zhou](#)

[pdfbibabs](#)

[Importance-based Neuron Allocation for Multilingual Neural Machine Translation](#)

[Wanying Xie](#) | [Yang Feng](#) | [Shuhao Gu](#) | [Dong Yu](#)

[pdfbibabs](#)

[Transfer Learning for Sequence Generation: from Single-source to Multi-source](#)

[Xuancheng Huang](#) | [Jingfang Xu](#) | [Maosong Sun](#) | [Yang Liu](#)

[pdfbibabs](#)

[A Closer Look at Few-Shot Crosslingual Transfer: The Choice of Shots Matters](#)

[Mengjie Zhao](#) | [Yi Zhu](#) | [Ehsan Shareghi](#) | [Ivan Vulić](#) | [Roi Reichart](#) | [Anna Korhonen](#) | [Hinrich Schütze](#)

[pdfbibabs](#)

[Coreference Reasoning in Machine Reading Comprehension](#)

[Mingzhu Wu](#) | [Nafise Sadat Moosavi](#) | [Dan Roth](#) | [Iryna Gurevych](#)

[pdfbibabs](#)

[Adapting Unsupervised Syntactic Parsing Methodology for Discourse Dependency Parsing](#)

[Liwen Zhang](#) | [Ge Wang](#) | [Wenjuan Han](#) | [Kewei Tu](#)

[pdfbibabs](#)

[A Conditional Splitting Framework for Efficient Constituency Parsing](#)

[Thanh-Tung Nguyen](#) | [Xuan-Phi Nguyen](#) | [Shafiq Joty](#) | [Xiaoli Li](#)

[pdfbibabs](#)

[A Unified Generative Framework for Various NER Subtasks](#)

[Hang Yan](#) | [Tao Gui](#) | [Junqi Dai](#) | [Qipeng Guo](#) | [Zheng Zhang](#) | [Xipeng Qiu](#)

[pdfbibabs](#)

[An In-depth Study on Internal Structure of Chinese Words](#)

[Chen Gong](#) | [Saihao Huang](#) | [Houquan Zhou](#) | [Zhenghua Li](#) | [Min Zhang](#) | [Zhefeng Wang](#) | [Baoxing Huai](#) | [Nicholas Jing Yuan](#)

[pdfbibabs](#)

[MulDA: A Multilingual Data Augmentation Framework for Low-Resource Cross-Lingual NER](#)

[Linlin Liu](#) | [Bosheng Ding](#) | [Lidong Bing](#) | [Shafiq Joty](#) | [Luo Si](#) | [Chunyan Miao](#)

[pdfbibabs](#)

[Lexicon Enhanced Chinese Sequence Labeling Using BERT Adapter](#)

[Wei Liu](#) | [Xiyan Fu](#) | [Yue Zhang](#) | [Wenming Xiao](#)

[pdfbibabs](#)

[Math Word Problem Solving with Explicit Numerical Values](#)

[Qinzhuo Wu](#) | [Qi Zhang](#) | [Zhongyu Wei](#) | [Xuanjing Huang](#)

[pdfbibabs](#)

[Neural-Symbolic Solver for Math Word Problems with Auxiliary Tasks](#)

[Jinghui Qin](#) | [Xiaodan Liang](#) | [Yining Hong](#) | [Jianheng Tang](#) | [Liang Lin](#)

[pdfbibabs](#)

[SMedBERT: A Knowledge-Enhanced Pre-trained Language Model with Structured Semantics for Medical Text Mining](#)

[Taolin Zhang](#) | [Zerui Cai](#) | [Chengyu Wang](#) | [Minghui Qiu](#) | [Bite Yang](#) | [Xiaofeng He](#)

[pdfbibabs](#)

[What is Your Article Based On? Inferring Fine-grained Provenance](#)

[Yi Zhang](#) | [Zachary Ives](#) | [Dan Roth](#)

[pdfbibabs](#)

[Cross-modal Memory Networks for Radiology Report Generation](#)

[Zhihong Chen](#) | [Yaling Shen](#) | [Yan Song](#) | [Xiang Wan](#)

[pdfbibabs](#)

[Controversy and Conformity: from Generalized to Personalized Aggressiveness Detection](#)

[Kamil Kanclerz](#) | [Alicja Figas](#) | [Marcin Gruza](#) | [Tomasz Kajdanowicz](#) | [Jan Kocon](#) | [Daria Puchalska](#)
| [Przemyslaw Kazienko](#)

[pdfbibabs](#)

[Multi-perspective Coherent Reasoning for Helpfulness Prediction of Multimodal Reviews](#)

[Junhao Liu](#) | [Zhen Hai](#) | [Min Yang](#) | [Lidong Bing](#)

[pdfbibabs](#)

[Instantaneous Grammatical Error Correction with Shallow Aggressive Decoding](#)

[Xin Sun](#) | [Tao Ge](#) | [Furu Wei](#) | [Houfeng Wang](#)

[pdfbibabs](#)

[Automatic ICD Coding via Interactive Shared Representation Networks with Self-distillation Mechanism](#)

[Tong Zhou](#) | [Pengfei Cao](#) | [Yubo Chen](#) | [Kang Liu](#) | [Jun Zhao](#) | [Kun Niu](#) | [Weifeng Chong](#) | [Shengping Liu](#)

[pdfbibabs](#)

[PHMOSpell: Phonological and Morphological Knowledge Guided Chinese Spelling Check](#)

[Li Huang](#) | [Junjie Li](#) | [Weiwei Jiang](#) | [Zhiyu Zhang](#) | [Minchuan Chen](#) | [Shaojun Wang](#) | [Jing Xiao](#)

[pdfbibabs](#)

[Guiding the Growth: Difficulty-Controllable Question Generation through Step-by-Step Rewriting](#)

[Yi Cheng](#) | [Siyao Li](#) | [Bang Liu](#) | [Ruihui Zhao](#) | [Sujian Li](#) | [Chenghua Lin](#) | [Yefeng Zheng](#)

[pdfbibabs](#)

[Improving Encoder by Auxiliary Supervision Tasks for Table-to-Text Generation](#)

[Liang Li](#) | [Can Ma](#) | [Yinliang Yue](#) | [Dayong Hu](#)

[pdfbibabs](#)

[POS-Constrained Parallel Decoding for Non-autoregressive Generation](#)

[Kexin Yang](#) | [Wenqiang Lei](#) | [Dayiheng Liu](#) | [Weizhen Qi](#) | [Jiancheng Lv](#)

[pdfbibabs](#)

[Bridging Subword Gaps in Pretrain-Finetune Paradigm for Natural Language Generation](#)

[Xin Liu](#) | [Baosong Yang](#) | [Dayiheng Liu](#) | [Haibo Zhang](#) | [Weihua Luo](#) | [Min Zhang](#) | [Haiying Zhang](#)
| [Jinsong Su](#)

[pdfbibabs](#)

[TGEA: An Error-Annotated Dataset and Benchmark Tasks for Text Generation from Pretrained Language Models](#)

[Jie He](#) | [Bo Peng](#) | [Yi Liao](#) | [Qun Liu](#) | [Deyi Xiong](#)

[pdfbibabs](#)

[Long-Span Summarization via Local Attention and Content Selection](#)

[Potsawee Manakul](#) | [Mark Gales](#)

[pdfbibabs](#)

[RepSum: Unsupervised Dialogue Summarization based on Replacement Strategy](#)

[Xiyan Fu](#) | [Yating Zhang](#) | [Tianyi Wang](#) | [Xiaozhong Liu](#) | [Changlong Sun](#) | [Zhenglu Yang](#)

[pdfbibabs](#)

[BASS: Boosting Abstractive Summarization with Unified Semantic Graph](#)

[Wenhao Wu](#) | [Wei Li](#) | [Xinyan Xiao](#) | [Jiachen Liu](#) | [Ziqiang Cao](#) | [Sujian Li](#) | [Hua Wu](#) | [Haifeng Wang](#)

[pdfbibabs](#)

[Capturing Relations between Scientific Papers: An Abstractive Model for Related Work Section Generation](#)

[Xiuying Chen](#) | [Hind Alamo](#) | [Mingzhe Li](#) | [Shen Gao](#) | [Xiangliang Zhang](#) | [Dongyan Zhao](#) | [Rui Yan](#)

[pdfbibabs](#)

[Focus Attention: Promoting Faithfulness and Diversity in Summarization](#)

[Rahul Aralikatte](#) | [Shashi Narayan](#) | [Joshua Maynez](#) | [Sascha Rothe](#) | [Ryan McDonald](#)

[pdfbibabs](#)

[Generating Query Focused Summaries from Query-Free Resources](#)

[Yumo Xu](#) | [Mirella Lapata](#)

[pdfbibabs](#)

[Robustifying Multi-hop QA through Pseudo-Evidentiality Training](#)

[Kyungjae Lee](#) | [Seung-won Hwang](#) | [Sang-eun Han](#) | [Dohyeon Lee](#)

[pdfbibabs](#)

[xMoCo: Cross Momentum Contrastive Learning for Open-Domain Question Answering](#)

[Nan Yang](#) | [Furu Wei](#) | [Binxing Jiao](#) | [Daxing Jiang](#) | [Linjun Yang](#)

[pdfbibabs](#)

[Learn to Resolve Conversational Dependency: A Consistency Training Framework for Conversational Question Answering](#)

[Gangwoo Kim](#) | [Hyunjae Kim](#) | [Jungsoo Park](#) | [Jaewoo Kang](#)

[pdfbibabs](#)

[PhotoChat: A Human-Human Dialogue Dataset With Photo Sharing Behavior For Joint Image-Text Modeling](#)

[Xiaoxue Zang](#) | [Lijuan Liu](#) | [Maria Wang](#) | [Yang Song](#) | [Hao Zhang](#) | [Jindong Chen](#)

[pdfbibabs](#)

Good for Misconceived Reasons: An Empirical Revisiting on the Need for Visual Context in Multimodal Machine Translation

[Zhiyong Wu](#) | [Lingpeng Kong](#) | [Wei Bi](#) | [Xiang Li](#) | [Ben Kao](#)

[pdfbibabs](#)

Attend What You Need: Motion-Appearance Synergistic Networks for Video Question Answering

[Ahjeong Seo](#) | [Gi-Cheon Kang](#) | [Joonhan Park](#) | [Byoung-Tak Zhang](#)

[pdfbibabs](#)

BERTifying the Hidden Markov Model for Multi-Source Weakly Supervised Named Entity Recognition

[Yinghao Li](#) | [Pranav Shetty](#) | [Lucas Liu](#) | [Chao Zhang](#) | [Le Song](#)

[pdfbibabs](#)

CIL: Contrastive Instance Learning Framework for Distantly Supervised Relation Extraction

[Tao Chen](#) | [Haizhou Shi](#) | [Siliang Tang](#) | [Zhigang Chen](#) | [Fei Wu](#) | [Yueting Zhuang](#)

[pdfbibabs](#)

SENT: Sentence-level Distant Relation Extraction via Negative Training

[Ruotian Ma](#) | [Tao Gui](#) | [Linyang Li](#) | [Qi Zhang](#) | [Xuanjing Huang](#) | [Yaqian Zhou](#)

[pdfbibabs](#)

An End-to-End Progressive Multi-Task Learning Framework for Medical Named Entity Recognition and Normalization

[Baohang Zhou](#) | [Xiangrui Cai](#) | [Ying Zhang](#) | [Xiaojie Yuan](#)

[pdfbibabs](#)

PRGC: Potential Relation and Global Correspondence Based Joint Relational Triple Extraction

[Hengyi Zheng](#) | [Rui Wen](#) | [Xi Chen](#) | [Yifan Yang](#) | [Yunyan Zhang](#) | [Ziheng Zhang](#) | [Ningyu Zhang](#) | [Bin Qin](#) | [Xu Ming](#) | [Yefeng Zheng](#)

[pdfbibabs](#)

Learning from Miscellaneous Other-Class Words for Few-shot Named Entity Recognition

[Meihan Tong](#) | [Shuai Wang](#) | [Bin Xu](#) | [Yixin Cao](#) | [Minghui Liu](#) | [Lei Hou](#) | [Juanzi Li](#)

[pdfbibabs](#)

Joint Biomedical Entity and Relation Extraction with Knowledge-Enhanced Collective Inference

[Tuan Lai](#) | [Heng Ji](#) | [ChengXiang Zhai](#) | [Quan Hung Tran](#)

[pdfbibabs](#)

Fine-grained Information Extraction from Biomedical Literature based on Knowledge-enriched Abstract Meaning Representation

[Zixuan Zhang](#) | [Nikolaus Parulian](#) | [Heng Ji](#) | [Ahmed Elsayed](#) | [Skatje Myers](#) | [Martha Palmer](#)

[pdfbibabs](#)

Unleash GPT-2 Power for Event Detection

[Amir Pouran Ben Veyseh](#) | [Viet Lai](#) | [Franck Dernoncourt](#) | [Thien Huu Nguyen](#)

[pdfbibabs](#)

CLEVE: Contrastive Pre-training for Event Extraction

[Ziqi Wang](#) | [Xiaozhi Wang](#) | [Xu Han](#) | [Yankai Lin](#) | [Lei Hou](#) | [Zhiyuan Liu](#) | [Peng Li](#) | [Juanzi Li](#) | [Jie Zhou](#)

[pdfbibabs](#)

[Document-level Event Extraction via Parallel Prediction Networks](#)

[Hang Yang](#) | [Dianbo Sui](#) | [Yubo Chen](#) | [Kang Liu](#) | [Jun Zhao](#) | [Taifeng Wang](#)

[pdfbibabs](#)

[StructuralLM: Structural Pre-training for Form Understanding](#)

[Chenliang Li](#) | [Bin Bi](#) | [Ming Yan](#) | [Wei Wang](#) | [Songfang Huang](#) | [Fei Huang](#) | [Luo Si](#)

[pdfbibabs](#)

[Dual Graph Convolutional Networks for Aspect-based Sentiment Analysis](#)

[Ruifan Li](#) | [Hao Chen](#) | [Fangxiang Feng](#) | [Zhanyu Ma](#) | [Xiaojie Wang](#) | [Eduard Hovy](#)

[pdfbibabs](#)

[Multi-Label Few-Shot Learning for Aspect Category Detection](#)

[Mengting Hu](#) | [Shiwan Zhao](#) | [Honglei Guo](#) | [Chao Xue](#) | [Hang Gao](#) | [Tiegang Gao](#) | [Renhong Cheng](#) | [Zhong Su](#)

[pdfbibabs](#)

[Argument Pair Extraction via Attention-guided Multi-Layer Multi-Cross Encoding](#)

[Liyang Cheng](#) | [Tianyu Wu](#) | [Lidong Bing](#) | [Luo Si](#)

[pdfbibabs](#)

[A Neural Transition-based Model for Argumentation Mining](#)

[Jianzhu Bao](#) | [Chuang Fan](#) | [Jipeng Wu](#) | [Yixue Dang](#) | [Jiachen Du](#) | [Ruifeng Xu](#)

[pdfbibabs](#)

[Keep It Simple: Unsupervised Simplification of Multi-Paragraph Text](#)

[Philippe Laban](#) | [Tobias Schnabel](#) | [Paul Bennett](#) | [Marti A. Hearst](#)

[pdfbibabs](#)

[Long Text Generation by Modeling Sentence-Level and Discourse-Level Coherence](#)

[Jian Guan](#) | [Xiaoxi Mao](#) | [Changjie Fan](#) | [Zitao Liu](#) | [Wenbiao Ding](#) | [Minlie Huang](#)

[pdfbibabs](#)

[OpenMEVA: A Benchmark for Evaluating Open-ended Story Generation Metrics](#)

[Jian Guan](#) | [Zhexin Zhang](#) | [Zhuoer Feng](#) | [Zitao Liu](#) | [Wenbiao Ding](#) | [Xiaoxi Mao](#) | [Changjie Fan](#) | [Minlie Huang](#)

[pdfbibabs](#)

[DYPLOC: Dynamic Planning of Content Using Mixed Language Models for Text Generation](#)

[Xinyu Hua](#) | [Ashwin Sreevatsa](#) | [Lu Wang](#)

[pdfbibabs](#)

[Controllable Open-ended Question Generation with A New Question Type Ontology](#)

[Shuyang Cao](#) | [Lu Wang](#)

[pdfbibabs](#)

[BERTGen: Multi-task Generation through BERT](#)

[Faidon Mitzalis](#) | [Ozan Caglayan](#) | [Pranava Madhyastha](#) | [Lucia Specia](#)

[pdfbibabs](#)

[Selective Knowledge Distillation for Neural Machine Translation](#)

[Fusheng Wang](#) | [Jianhao Yan](#) | [Fandong Meng](#) | [Jie Zhou](#)

[pdfbibabs](#)

[Measuring and Increasing Context Usage in Context-Aware Machine Translation](#)

[Patrick Fernandes](#) | [Kayo Yin](#) | [Graham Neubig](#) | [André F. T. Martins](#)

[pdfbibabs](#)

[Beyond Offline Mapping: Learning Cross-lingual Word Embeddings through Context Anchoring](#)

[Aitor Ormazabal](#) | [Mikel Artetxe](#) | [Aitor Soroa](#) | [Gorka Labaka](#) | [Eneko Agirre](#)

[pdfbibabs](#)

[CCMatrix: Mining Billions of High-Quality Parallel Sentences on the Web](#)

[Holger Schwenk](#) | [Guillaume Wenzek](#) | [Sergey Edunov](#) | [Edouard Grave](#) | [Armand Joulin](#) | [Angela Fan](#)

[pdfbibabs](#)

[Length-Adaptive Transformer: Train Once with Length Drop, Use Anytime with Search](#)

[Gyuwan Kim](#) | [Kyunghyun Cho](#)

[pdfbibabs](#)

[GhostBERT: Generate More Features with Cheap Operations for BERT](#)

[Zhiqi Huang](#) | [Lu Hou](#) | [Lifeng Shang](#) | [Xin Jiang](#) | [Xiao Chen](#) | [Qun Liu](#)

[pdfbibabs](#)

[Super Tickets in Pre-Trained Language Models: From Model Compression to Improving Generalization](#)

[Chen Liang](#) | [Simiao Zuo](#) | [Minshuo Chen](#) | [Haoming Jiang](#) | [Xiaodong Liu](#) | [Pengcheng He](#) | [Tuo Zhao](#) | [Weizhu Chen](#)

[pdfbibabs](#)

[A Novel Estimator of Mutual Information for Learning to Disentangle Textual Representations](#)

[Pierre Colombo](#) | [Pablo Piantanida](#) | [Chloé Clavel](#)

[pdfbibabs](#)

[Determinantal Beam Search](#)

[Clara Meister](#) | [Martina Forster](#) | [Ryan Cotterell](#)

[pdfbibabs](#)

[Multi-hop Graph Convolutional Network with High-order Chebyshev Approximation for Text Reasoning](#)

[Shuoran Jiang](#) | [Qingcai Chen](#) | [Xin Liu](#) | [Baotian Hu](#) | [Lisai Zhang](#)

[pdfbibabs](#)

[Accelerating Text Communication via Abbreviated Sentence Input](#)

[Jiban Adhikary](#) | [Jamie Berger](#) | [Keith Vertanen](#)

[pdfbibabs](#)

[Regression Bugs Are In Your Model! Measuring, Reducing and Analyzing Regressions In NLP Model Updates](#)

[Yuqing Xie](#) | [Yi-An Lai](#) | [Yuanjun Xiong](#) | [Yi Zhang](#) | [Stefano Soatto](#)

[pdfbibabs](#)

[Detecting Propaganda Techniques in Memes](#)

[Dimitar Dimitrov](#) | [Bishr Bin Ali](#) | [Shaden Shaar](#) | [Firoj Alam](#) | [Fabrizio Silvestri](#) | [Hamed Firooz](#) | [Preslav Nakov](#) | [Giovanni Da San Martino](#)

[pdfbibabs](#)

[On the Efficacy of Adversarial Data Collection for Question Answering: Results from a Large-Scale Randomized Study](#)

[Divyansh Kaushik](#) | [Douwe Kiela](#) | [Zachary C. Lipton](#) | [Wen-tau Yih](#)

[pdfbibabs](#)

Learning Dense Representations of Phrases at Scale

[Jinhyuk Lee](#) | [Mujeen Sung](#) | [Jaewoo Kang](#) | [Danqi Chen](#)

[pdfbibabs](#)

End-to-End Training of Neural Retrievers for Open-Domain Question Answering

[Devendra Sachan](#) | [Mostofa Patwary](#) | [Mohammad Shoeybi](#) | [Neel Kant](#) | [Wei Ping](#) | [William L. Hamilton](#) | [Bryan Catanzaro](#)

[pdfbibabs](#)

Question Answering Over Temporal Knowledge Graphs

[Apoorv Saxena](#) | [Soumen Chakrabarti](#) | [Partha Talukdar](#)

[pdfbibabs](#)

Language Model Augmented Relevance Score

[Ruibo Liu](#) | [Jason Wei](#) | [Soroush Vosoughi](#)

[pdfbibabs](#)

DExperts: Decoding-Time Controlled Text Generation with Experts and Anti-Experts

[Alisa Liu](#) | [Maarten Sap](#) | [Ximing Lu](#) | [Swabha Swayamdipta](#) | [Chandra Bhagavatula](#) | [Noah A. Smith](#) | [Yejin Choi](#)

[pdfbibabs](#)

Polyjuice: Generating Counterfactuals for Explaining, Evaluating, and Improving Models

[Tongshuang Wu](#) | [Marco Tulio Ribeiro](#) | [Jeffrey Heer](#) | [Daniel Weld](#)

[pdfbibabs](#)

Metaphor Generation with Conceptual Mappings

[Kevin Stowe](#) | [Tuhin Chakrabarty](#) | [Nanyun Peng](#) | [Smaranda Muresan](#) | [Iryna Gurevych](#)

[pdfbibabs](#)

Learning Latent Structures for Cross Action Phrase Relations in Wet Lab Protocols

[Chaitanya Kulkarni](#) | [Jany Chan](#) | [Eric Fosler-Lussier](#) | [Raghu Machiraju](#)

[pdfbibabs](#)

Multimodal Multi-Speaker Merger & Acquisition Financial Modeling: A New Task, Dataset, and Neural Baselines

[Ramit Sawhney](#) | [Mihir Goyal](#) | [Prakhar Goel](#) | [Puneet Mathur](#) | [Rajiv Ratn Shah](#)

[pdfbibabs](#)

Mid-Air Hand Gestures for Post-Editing of Machine Translation

[Rashad Albo Jamara](#) | [Nico Herbig](#) | [Antonio Krüger](#) | [Josef van Genabith](#)

[pdfbibabs](#)

Inter-GPS: Interpretable Geometry Problem Solving with Formal Language and Symbolic Reasoning

[Pan Lu](#) | [Ran Gong](#) | [Shibiao Jiang](#) | [Liang Qiu](#) | [Siyuan Huang](#) | [Xiaodan Liang](#) | [Song-Chun Zhu](#)

[pdfbibabs](#)

Joint Verification and Reranking for Open Fact Checking Over Tables

[Michael Sejr Schlichtkrull](#) | [Vladimir Karpukhin](#) | [Barlas Oguz](#) | [Mike Lewis](#) | [Wen-tau Yih](#) | [Sebastian Riedel](#)

[pdfbibabs](#)

Evaluation of Thematic Coherence in Microblogs

[Iman Munire Bilal](#) | [Bo Wang](#) | [Maria Liakata](#) | [Rob Procter](#) | [Adam Tsakalidis](#)

[pdfbibabs](#)

[Neural semi-Markov CRF for Monolingual Word Alignment](#)

[Wuwei Lan](#) | [Chao Jiang](#) | [Wei Xu](#)

[pdfbibabs](#)

[Privacy at Scale: Introducing the PrivaSeer Corpus of Web Privacy Policies](#)

[Mukund Srinath](#) | [Shomir Wilson](#) | [C Lee Giles](#)

[pdfbibabs](#)

[The statistical advantage of automatic NLG metrics at the system level](#)

[Johnny Wei](#) | [Robin Jia](#)

[pdfbibabs](#)

[Are Missing Links Predictable? An Inferential Benchmark for Knowledge Graph Completion](#)

[Yixin Cao](#) | [Xiang Ji](#) | [Xin Lv](#) | [Juanzi Li](#) | [Yonggang Wen](#) | [Hanwang Zhang](#)

[pdfbibabs](#)

[ConvoSumm: Conversation Summarization Benchmark and Improved Abstractive Summarization with Argument Mining](#)

[Alexander Fabbri](#) | [Faiaz Rahman](#) | [Imad Rizvi](#) | [Borui Wang](#) | [Haoran Li](#) | [Yashar Mehdad](#) | [Dragomir Radev](#)

[pdfbibabs](#)

[Improving Factual Consistency of Abstractive Summarization via Question Answering](#)

[Feng Nan](#) | [Cicero Nogueira dos Santos](#) | [Henghui Zhu](#) | [Patrick Ng](#) | [Kathleen McKeown](#) | [Ramesh Nallapati](#) | [Dejiao Zhang](#) | [Zhiguo Wang](#) | [Andrew O. Arnold](#) | [Bing Xiang](#)

[pdfbibabs](#)

[EmailSum: Abstractive Email Thread Summarization](#)

[Shiyue Zhang](#) | [Asli Celikyilmaz](#) | [Jianfeng Gao](#) | [Mohit Bansal](#)

[pdfbibabs](#)

[Cross-Lingual Abstractive Summarization with Limited Parallel Resources](#)

[Yu Bai](#) | [Yang Gao](#) | [Heyan Huang](#)

[pdfbibabs](#)

[Dissecting Generation Modes for Abstractive Summarization Models via Ablation and Attribution](#)

[Jiacheng Xu](#) | [Greg Durrett](#)

[pdfbibabs](#)

[Learning Prototypical Functions for Physical Artifacts](#)

[Tianyu Jiang](#) | [Ellen Riloff](#)

[pdfbibabs](#)

[Verb Knowledge Injection for Multilingual Event Processing](#)

[Olga Majewska](#) | [Ivan Vulić](#) | [Goran Glavaš](#) | [Edoardo Maria Ponti](#) | [Anna Korhonen](#)

[pdfbibabs](#)

[Dynamic Contextualized Word Embeddings](#)

[Valentin Hofmann](#) | [Janet Pierrehumbert](#) | [Hinrich Schütze](#)

[pdfbibabs](#)

[Lexical Semantic Change Discovery](#)

[Sinan Kurtçiyigit](#) | [Maike Park](#) | [Dominik Schlechtweg](#) | [Jonas Kuhn](#) | [Sabine Schulte im Walde](#)

[pdfbibabs](#)

The R-U-A-Robot Dataset: Helping Avoid Chatbot Deception by Detecting User Questions About Human or Non-Human Identity

[David Gros](#) | [Yu Li](#) | [Zhou Yu](#)

[pdfbibabs](#)

Using Meta-Knowledge Mined from Identifiers to Improve Intent Recognition in Conversational Systems

[Claudio Pinhanez](#) | [Paulo Cavalin](#) | [Victor Henrique Alves Ribeiro](#) | [Ana Appel](#) | [Heloisa Candello](#) | [Julio Nogima](#) | [Mauro Pichiliani](#) | [Melina Guerra](#) | [Maira de Bayser](#) | [Gabriel Malfatti](#) | [Henrique Ferreira](#)

[pdfbibabs](#)

Space Efficient Context Encoding for Non-Task-Oriented Dialogue Generation with Graph Attention Transformer

[Fabian Galetzka](#) | [Jewgeni Rose](#) | [David Schlangen](#) | [Jens Lehmann](#)

[pdfbibabs](#)

DialogueCRN: Contextual Reasoning Networks for Emotion Recognition in Conversations

[Dou Hu](#) | [Lingwei Wei](#) | [Xiaoyong Huai](#)

[pdfbibabs](#)

Cross-replication Reliability - An Empirical Approach to Interpreting Inter-rater Reliability

[Ka Wong](#) | [Praveen Paritosh](#) | [Lora Aroyo](#)

[pdfbibabs](#)

TIMEDIAL: Temporal Commonsense Reasoning in Dialog

[Lianhui Qin](#) | [Aditya Gupta](#) | [Shyam Upadhyay](#) | [Luheng He](#) | [Yejin Choi](#) | [Manaal Faruqui](#)

[pdfbibabs](#)

RAW-C: Relatedness of Ambiguous Words in Context (A New Lexical Resource for English)

[Sean Trott](#) | [Benjamin Bergen](#)

[pdfbibabs](#)

ARBERT & MARBERT: Deep Bidirectional Transformers for Arabic

[Muhammad Abdul-Mageed](#) | [AbdelRahim Elmadany](#) | [El Moatez Billah Nagoudi](#)

[pdfbibabs](#)

Improving Paraphrase Detection with the Adversarial Paraphrasing Task

[Animesh Nighojkar](#) | [John Licato](#)

[pdfbibabs](#)

ADEPT: An Adjective-Dependent Plausibility Task

[Ali Emami](#) | [Ian Porada](#) | [Alexandra Olteanu](#) | [Kaheer Suleman](#) | [Adam Trischler](#) | [Jackie Chi Kit Cheung](#)

[pdfbibabs](#)

ReadOnce Transformers: Reusable Representations of Text for Transformers

[Shih-Ting Lin](#) | [Ashish Sabharwal](#) | [Tushar Khot](#)

[pdfbibabs](#)

Conditional Generation of Temporally-ordered Event Sequences

[Shih-Ting Lin](#) | [Nathanael Chambers](#) | [Greg Durrett](#)

[pdfbibabs](#)

Hate Speech Detection Based on Sentiment Knowledge Sharing

[Xianbing Zhou](#) | [Yang Yong](#) | [Xiaochao Fan](#) | [Ge Ren](#) | [Yunfeng Song](#) | [Yufeng Diao](#) | [Liang Yang](#) | [Hongfei Lin](#)

[pdfbibabs](#)

Transition-based Bubble Parsing: Improvements on Coordination Structure Prediction

[Tianze Shi](#) | [Lillian Lee](#)

[pdfbibabs](#)

SpanNER: Named Entity Re-/Recognition as Span Prediction

[Jinlan Fu](#) | [Xuanjing Huang](#) | [Pengfei Liu](#)

[pdfbibabs](#)

StructFormer: Joint Unsupervised Induction of Dependency and Constituency Structure from Masked Language Modeling

[Yikang Shen](#) | [Yi Tay](#) | [Che Zheng](#) | [Dara Bahri](#) | [Donald Metzler](#) | [Aaron Courville](#)

[pdfbibabs](#)

Language Embeddings for Typology and Cross-lingual Transfer Learning

[Dian Yu](#) | [Taiqi He](#) | [Kenji Sagae](#)

[pdfbibabs](#)

Can Sequence-to-Sequence Models Crack Substitution Ciphers?

[Nada Aldarrab](#) | [Jonathan May](#)

[pdfbibabs](#)

Beyond Noise: Mitigating the Impact of Fine-grained Semantic Divergences on Neural Machine Translation

[Eleftheria Briakou](#) | [Marine Carpuat](#)

[pdfbibabs](#)

Discriminative Reranking for Neural Machine Translation

[Ann Lee](#) | [Michael Auli](#) | [Marc'Aurelio Ranzato](#)

[pdfbibabs](#)

Mind Your Outliers! Investigating the Negative Impact of Outliers on Active Learning for Visual Question Answering

[Siddharth Karamcheti](#) | [Ranjay Krishna](#) | [Li Fei-Fei](#) | [Christopher Manning](#)

[pdfbibabs](#)

All That's 'Human' Is Not Gold: Evaluating Human Evaluation of Generated Text

[Elizabeth Clark](#) | [Tal August](#) | [Sofia Serrano](#) | [Nikita Haduong](#) | [Suchin Gururangan](#) | [Noah A. Smith](#)

[pdfbibabs](#)

Scientific Credibility of Machine Translation Research: A Meta-Evaluation of 769 Papers

[Benjamin Marie](#) | [Atsushi Fujita](#) | [Raphael Rubino](#)

[pdfbibabs](#)

Neural Machine Translation with Monolingual Translation Memory

[Deng Cai](#) | [Yan Wang](#) | [Huayang Li](#) | [Wai Lam](#) | [Lemao Liu](#)

[pdfbibabs](#)

Intrinsic Dimensionality Explains the Effectiveness of Language Model Fine-Tuning

[Armen Aghajanyan](#) | [Sonal Gupta](#) | [Luke Zettlemoyer](#)

[pdfbibabs](#)

UnNatural Language Inference

[Koustuv Sinha](#) | [Prasanna Parthasarathi](#) | [Joelle Pineau](#) | [Adina Williams](#)

[pdfbibabs](#)

[Including Signed Languages in Natural Language Processing](#)

[Kayo Yin](#) | [Amit Moryossef](#) | [Julie Hochgesang](#) | [Yoav Goldberg](#) | [Malihe Alikhani](#)

[pdfbibabs](#)

[Vocabulary Learning via Optimal Transport for Neural Machine Translation](#)

[Jingjing Xu](#) | [Hao Zhou](#) | [Chun Gan](#) | [Zaixiang Zheng](#) | [Lei Li](#)