

- 1, TITLE: Knowledge Router: Learning Disentangled Representations for Knowledge Graphs  
<https://www.aclweb.org/anthology/2021.naacl-main.1>  
AUTHORS: Shuai Zhang, Xi Rao, Yi Tay, Ce Zhang  
HIGHLIGHT: To this end, this paper proposes to learn disentangled representations of KG entities - a new method that disentangles the inner latent properties of KG entities.
- 2, TITLE: Distantly Supervised Relation Extraction with Sentence Reconstruction and Knowledge Base Priors  
<https://www.aclweb.org/anthology/2021.naacl-main.2>  
AUTHORS: Fenia Christopoulou, Makoto Miwa, Sophia Ananiadou  
HIGHLIGHT: We propose a multi-task, probabilistic approach to facilitate distantly supervised relation extraction by bringing closer the representations of sentences that contain the same Knowledge Base pairs.
- 3, TITLE: Cross-Task Instance Representation Interactions and Label Dependencies for Joint Information Extraction with Graph Convolutional Networks  
<https://www.aclweb.org/anthology/2021.naacl-main.3>  
AUTHORS: Minh Van Nguyen, Viet Lai, Thien Huu Nguyen  
HIGHLIGHT: This paper presents a novel deep learning model to simultaneously solve the four tasks of IE in a single model (called FourIE).
- 4, TITLE: Abstract Meaning Representation Guided Graph Encoding and Decoding for Joint Information Extraction  
<https://www.aclweb.org/anthology/2021.naacl-main.4>  
AUTHORS: Zixuan Zhang, Heng Ji  
HIGHLIGHT: To take advantage of such similarity, we propose a novel AMR-guided framework for joint information extraction to discover entities, relations, and events with the help of a pre-trained AMR parser.
- 5, TITLE: A Frustratingly Easy Approach for Entity and Relation Extraction  
<https://www.aclweb.org/anthology/2021.naacl-main.5>  
AUTHORS: Zexuan Zhong, Danqi Chen  
HIGHLIGHT: In this work, we present a simple pipelined approach for entity and relation extraction, and establish the new state-of-the-art on standard benchmarks (ACE04, ACE05 and SciERC), obtaining a 1.7%-2.8% absolute improvement in relation F1 over previous joint models with the same pre-trained encoders.
- 6, TITLE: Event Time Extraction and Propagation via Graph Attention Networks  
<https://www.aclweb.org/anthology/2021.naacl-main.6>  
AUTHORS: Haoyang Wen, Yanru Qu, Heng Ji, Qiang Ning, Jiawei Han, Avi Sil, Hanghang Tong, Dan Roth  
HIGHLIGHT: This paper first formulates this problem based on a 4-tuple temporal representation used in entity slot filling, which allows us to represent fuzzy time spans more conveniently. We then propose a graph attention network-based approach to propagate temporal information over document-level event graphs constructed by shared entity arguments and temporal relations.
- 7, TITLE: Probing Word Translations in the Transformer and Trading Decoder for Encoder Layers  
<https://www.aclweb.org/anthology/2021.naacl-main.7>  
AUTHORS: Hongfei Xu, Josef van Genabith, Qiuhui Liu, Deyi Xiong  
HIGHLIGHT: In this work, we show that this is not quite the case: translation already happens progressively in encoder layers and even in the input embeddings.
- 8, TITLE: Mediators in Determining what Processing BERT Performs First  
<https://www.aclweb.org/anthology/2021.naacl-main.8>  
AUTHORS: Aviv Slobodkin, Leshem Choshen, Omri Abend  
HIGHLIGHT: As a test-case mediating factor, we consider the prediction's context length, namely the length of the span whose processing is minimally required to perform the prediction.
- 9, TITLE: Automatic Generation of Contrast Sets from Scene Graphs: Probing the Compositional Consistency of GQA  
<https://www.aclweb.org/anthology/2021.naacl-main.9>  
AUTHORS: Yonatan Bitton, Gabriel Stanovsky, Roy Schwartz, Michael Elhadad  
HIGHLIGHT: While most contrast sets were created manually, requiring intensive annotation effort, we present a novel method which leverages rich semantic input representation to automatically generate contrast sets for the visual question answering task.
- 10, TITLE: Multilingual Language Models Predict Human Reading Behavior  
<https://www.aclweb.org/anthology/2021.naacl-main.10>  
AUTHORS: Nora Hollenstein, Federico Pirovano, Ce Zhang, Lena J?ger, Lisa Beinborn

**HIGHLIGHT:** We analyze if large language models are able to predict patterns of human reading behavior.

11, **TITLE:** Do Syntactic Probes Probe Syntax? Experiments with Jabberwocky Probing

<https://www.aclweb.org/anthology/2021.naacl-main.11>

**AUTHORS:** Rowan Hall Maudslay, Ryan Cotterell

**HIGHLIGHT:** However, drawing a generalisation about a model's linguistic knowledge about a specific phenomena based on what a probe is able to learn may be problematic: in this work, we show that semantic cues in training data means that syntactic probes do not properly isolate syntax.

12, **TITLE:** A Non-Linear Structural Probe

<https://www.aclweb.org/anthology/2021.naacl-main.12>

**AUTHORS:** Jennifer C. White, Tiago Pimentel, Naomi Saphra, Ryan Cotterell

**HIGHLIGHT:** We examine the case of a structural probe (Hewitt and Manning, 2019), which aims to investigate the encoding of syntactic structure in contextual representations through learning only linear transformations.

13, **TITLE:** Concealed Data Poisoning Attacks on NLP Models

<https://www.aclweb.org/anthology/2021.naacl-main.13>

**AUTHORS:** Eric Wallace, Tony Zhao, Shi Feng, Sameer Singh

**HIGHLIGHT:** In this work, we develop a new data poisoning attack that allows an adversary to control model predictions whenever a desired trigger phrase is present in the input.

14, **TITLE:** Backtranslation Feedback Improves User Confidence in MT, Not Quality

<https://www.aclweb.org/anthology/2021.naacl-main.14>

**AUTHORS:** Vil'm Zouhar, Michal Nov?k, Mat?? ?ilinec, Ondrej Bojar, Mateo Obreg?n, Robin L. Hill, Fr?d?ric Blain, Marina Fomicheva, Lucia Specia, Lisa Yankovskaya

**HIGHLIGHT:** Translating text into a language unknown to the text's author, dubbed outbound translation, is a modern need for which the user experience has significant room for improvement, beyond the basic machine translation facility. We demonstrate this by showing three ways in which user confidence in the outbound translation, as well as its overall final quality, can be affected: backward translation, quality estimation (with alignment) and source paraphrasing.

15, **TITLE:** Data Filtering using Cross-Lingual Word Embeddings

<https://www.aclweb.org/anthology/2021.naacl-main.15>

**AUTHORS:** Christian Herold, Jan Rosendahl, Joris Vanvinckenroye, Hermann Ney

**HIGHLIGHT:** First, we analyze the performance of language identification, a tool commonly used for data filtering in the MT community and identify specific weaknesses. Based on our findings, we then propose several novel methods for data filtering, based on cross-lingual word embeddings.

16, **TITLE:** Improving the Lexical Ability of Pretrained Language Models for Unsupervised Neural Machine Translation

<https://www.aclweb.org/anthology/2021.naacl-main.16>

**AUTHORS:** Alexandra Chronopoulou, Dario Stojanovski, Alexander Fraser

**HIGHLIGHT:** In this paper, we enhance the bilingual masked language model pretraining with lexical-level information by using type-level cross-lingual subword embeddings.

17, **TITLE:** Neural Machine Translation without Embeddings

<https://www.aclweb.org/anthology/2021.naacl-main.17>

**AUTHORS:** Uri Shaham, Omer Levy

**HIGHLIGHT:** Surprisingly, replacing the ubiquitous embedding layer with one-hot representations of each byte does not hurt performance; experiments on byte-to-byte machine translation from English to 10 different languages show a consistent improvement in BLEU, rivaling character-level and even standard subword-level models.

18, **TITLE:** Counterfactual Data Augmentation for Neural Machine Translation

<https://www.aclweb.org/anthology/2021.naacl-main.18>

**AUTHORS:** Qi Liu, Matt Kusner, Phil Blunsom

**HIGHLIGHT:** We propose a data augmentation method for neural machine translation.

19, **TITLE:** Cultural and Geographical Influences on Image Translatability of Words across Languages

<https://www.aclweb.org/anthology/2021.naacl-main.19>

**AUTHORS:** Nikzad Khani, Isidora Tourni, Mohammad Sadegh Rasooli, Chris Callison-Burch, Derry Tanti Wijaya

**HIGHLIGHT:** To better understand when images are useful for translation, we study image translatability of words, which we define as the translatability of words via images, by measuring intra- and inter-cluster similarities of image representations of words that are translations of each other.

- 20, TITLE: Multilingual BERT Post-Pretraining Alignment  
<https://www.aclweb.org/anthology/2021.naacl-main.20>  
AUTHORS: Lin Pan, Chung-Wei Hang, Haode Qi, Abhishek Shah, Saloni Potdar, Mo Yu  
HIGHLIGHT: We propose a simple method to align multilingual contextual embeddings as a post-pretraining step for improved cross-lingual transferability of the pretrained language models.
- 21, TITLE: A Million Tweets Are Worth a Few Points: Tuning Transformers for Customer Service Tasks  
<https://www.aclweb.org/anthology/2021.naacl-main.21>  
AUTHORS: Amir Hadifar, Sofie Labat, Veronique Hoste, Chris Develder, Thomas Demeester  
HIGHLIGHT: We address this gap by collecting a multilingual social media corpus containing customer service conversations (865k tweets), comparing various pipelines of pretraining and finetuning approaches, applying them on 5 different end tasks.
- 22, TITLE: Paragraph-level Rationale Extraction through Regularization: A case study on European Court of Human Rights Cases  
<https://www.aclweb.org/anthology/2021.naacl-main.22>  
AUTHORS: Ilias Chalkidis, Manos Fergadiotis, Dimitrios Tsarapatsanis, Nikolaos Aletras, Ion Androutsopoulos, Prodrimos Malakasiotis  
HIGHLIGHT: To this end, we introduce a new application on legal text where, contrary to mainstream literature targeting word-level rationales, we conceive rationales as selected paragraphs in multi-paragraph structured court cases.
- 23, TITLE: Answering Product-Questions by Utilizing Questions from Other Contextually Similar Products  
<https://www.aclweb.org/anthology/2021.naacl-main.23>  
AUTHORS: Ohad Rozen, David Carmel, Avihai Mejer, Vitaly Mirkis, Yftah Ziser  
HIGHLIGHT: In this work, we propose a novel and complementary approach for predicting the answer for such questions, based on the answers for similar questions asked on similar products.
- 24, TITLE: EnSidNet: Enhanced Hybrid Siamese-Deep Network for grouping clinical trials into drug-development pathways  
<https://www.aclweb.org/anthology/2021.naacl-main.24>  
AUTHORS: Lucia Pagani  
HIGHLIGHT: Here we present an approach for the unmet need of drug-development pathway reconstruction, based on an Enhanced hybrid Siamese-Deep Neural Network (EnSidNet).
- 25, TITLE: DATE: Detecting Anomalies in Text via Self-Supervision of Transformers  
<https://www.aclweb.org/anthology/2021.naacl-main.25>  
AUTHORS: Andrei Manolache, Florin Brad, Elena Burceanu  
HIGHLIGHT: We use this approach for AD in text, by introducing a novel pretext task on text sequences.
- 26, TITLE: A Simple Approach for Handling Out-of-Vocabulary Identifiers in Deep Learning for Source Code  
<https://www.aclweb.org/anthology/2021.naacl-main.26>  
AUTHORS: Nadezhda Chirkova, Sergey Troshin  
HIGHLIGHT: We propose a simple, yet effective method, based on identifier anonymization, to handle out-of-vocabulary (OOV) identifiers.
- 27, TITLE: Fast and Scalable Dialogue State Tracking with Explicit Modular Decomposition  
<https://www.aclweb.org/anthology/2021.naacl-main.27>  
AUTHORS: Dingmin Wang, Chenghua Lin, Qi Liu, Kam-Fai Wong  
HIGHLIGHT: We present a fast and scalable architecture called Explicit Modular Decomposition (EMD), in which we incorporate both classification-based and extraction-based methods and design four modules (for classification and sequence labelling) to jointly extract dialogue states.
- 28, TITLE: Augmented SBERT: Data Augmentation Method for Improving Bi-Encoders for Pairwise Sentence Scoring Tasks  
<https://www.aclweb.org/anthology/2021.naacl-main.28>  
AUTHORS: Nandan Thakur, Nils Reimers, Johannes Daxenberger, Iryna Gurevych  
HIGHLIGHT: We present a simple yet efficient data augmentation strategy called Augmented SBERT, where we use the cross-encoder to label a larger set of input pairs to augment the training data for the bi-encoder.
- 29, TITLE: SmBoP: Semi-autoregressive Bottom-up Semantic Parsing

<https://www.aclweb.org/anthology/2021.naacl-main.29>

AUTHORS: Ohad Rubin, Jonathan Berant

HIGHLIGHT: In this work, we propose an alternative approach: a Semi-autoregressive Bottom-up Parser (SmBoP) that constructs at decoding step  $t$  the top- $K$  sub-trees of height  $= t$ .

30, TITLE: SGL: Speaking the Graph Languages of Semantic Parsing via Multilingual Translation

<https://www.aclweb.org/anthology/2021.naacl-main.30>

AUTHORS: Luigi Procopio, Rocco Tripodi, Roberto Navigli

HIGHLIGHT: In this work, instead, we reframe semantic parsing towards multiple formalisms as Multilingual Neural Machine Translation (MNMT), and propose SGL, a many-to-many seq2seq architecture trained with an MNMT objective.

31, TITLE: Unifying Cross-Lingual Semantic Role Labeling with Heterogeneous Linguistic Resources

<https://www.aclweb.org/anthology/2021.naacl-main.31>

AUTHORS: Simone Conia, Andrea Bacciu, Roberto Navigli

HIGHLIGHT: In this work, we address this issue and present a unified model to perform cross-lingual SRL over heterogeneous linguistic resources.

32, TITLE: Fool Me Twice: Entailment from Wikipedia Gamification

<https://www.aclweb.org/anthology/2021.naacl-main.32>

AUTHORS: Julian Eisenschlos, Bhuwan Dhingra, Jannis Bulian, Benjamin B?rschinger, Jordan Boyd-Graber

HIGHLIGHT: We release FoolMeTwice (FM2 for short), a large dataset of challenging entailment pairs collected through a fun multi-player game.

33, TITLE: Meta-Learning for Domain Generalization in Semantic Parsing

<https://www.aclweb.org/anthology/2021.naacl-main.33>

AUTHORS: Bailin Wang, Mirella Lapata, Ivan Titov

HIGHLIGHT: In this work, we use a meta-learning framework which targets zero-shot domain generalization for semantic parsing.

34, TITLE: Aspect-Controlled Neural Argument Generation

<https://www.aclweb.org/anthology/2021.naacl-main.34>

AUTHORS: Benjamin Schiller, Johannes Daxenberger, Iryna Gurevych

HIGHLIGHT: In this work, we present the Arg-CTRL - a language model for argument generation that can be controlled to generate sentence-level arguments for a given topic, stance, and aspect.

35, TITLE: Text Generation from Discourse Representation Structures

<https://www.aclweb.org/anthology/2021.naacl-main.35>

AUTHORS: Jiangming Liu, Shay B. Cohen, Mirella Lapata

HIGHLIGHT: We propose neural models to generate text from formal meaning representations based on Discourse Representation Structures (DRSs).

36, TITLE: APo-VAE: Text Generation in Hyperbolic Space

<https://www.aclweb.org/anthology/2021.naacl-main.36>

AUTHORS: Shuyang Dai, Zhe Gan, Yu Cheng, Chenyang Tao, Lawrence Carin, Jingjing Liu

HIGHLIGHT: In this paper, we investigate text generation in a hyperbolic latent space to learn continuous hierarchical representations.

37, TITLE: DART: Open-Domain Structured Data Record to Text Generation

<https://www.aclweb.org/anthology/2021.naacl-main.37>

AUTHORS: Linyong Nan, Dragomir Radev, Rui Zhang, Amrit Rau, Abhinand Sivaprasad, Chiachun Hsieh, Xiangru Tang, Aadit Vyas, Neha Verma, Pranav Krishna, Yangxiaokang Liu, Nadia Irwanto, Jessica Pan, Faiaz Rahman, Ahmad Zaidi, Mutethia Mutuma, Yasin Tarabar, Ankit Gupta, Tao Yu, Yi Chern Tan, Xi Victoria Lin, Caiming Xiong, Richard Socher, Nazneen Fatema Rajani

HIGHLIGHT: We present DART, an open domain structured DATA Record to Text generation dataset with over 82k instances (DARTs).

38, TITLE: When Being Unseen from mBERT is just the Beginning: Handling New Languages With Multilingual Language Models

<https://www.aclweb.org/anthology/2021.naacl-main.38>

AUTHORS: Benjamin Muller, Antonios Anastasopoulos, Beno?t Sagot, Djam? Seddah

**HIGHLIGHT:** In this work, by comparing multilingual and monolingual models, we show that such models behave in multiple ways on unseen languages.

39, **TITLE:** Multi-Adversarial Learning for Cross-Lingual Word Embeddings

<https://www.aclweb.org/anthology/2021.naacl-main.39>

**AUTHORS:** Haozhou Wang, James Henderson, Paola Merlo

**HIGHLIGHT:** We assume instead that, especially across distant languages, the mapping is only piece-wise linear, and propose a multi-adversarial learning method.

40, **TITLE:** Multi-view Subword Regularization

<https://www.aclweb.org/anthology/2021.naacl-main.40>

**AUTHORS:** Xinyi Wang, Sebastian Ruder, Graham Neubig

**HIGHLIGHT:** However, standard heuristic algorithms often lead to sub-optimal segmentation, especially for languages with limited amounts of data. In this paper, we take two major steps towards alleviating this problem.

41, **TITLE:** mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer

<https://www.aclweb.org/anthology/2021.naacl-main.41>

**AUTHORS:** Linting Xue, Noah Constant, Adam Roberts, Mihir Kale, Rami Al-Rfou, Aditya Siddhant, Aditya Barua, Colin Raffel

**HIGHLIGHT:** In this paper, we introduce mT5, a multilingual variant of T5 that was pre-trained on a new Common Crawl-based dataset covering 101 languages.

42, **TITLE:** MetaXL: Meta Representation Transformation for Low-resource Cross-lingual Learning

<https://www.aclweb.org/anthology/2021.naacl-main.42>

**AUTHORS:** Mengzhou Xia, Guoqing Zheng, Subhabrata Mukherjee, Milad Shokouhi, Graham Neubig, Ahmed Hassan Awadallah

**HIGHLIGHT:** In this paper, we propose MetaXL, a meta-learning based framework that learns to transform representations judiciously from auxiliary languages to a target one and brings their representation spaces closer for effective transfer.

43, **TITLE:** Open Domain Question Answering over Tables via Dense Retrieval

<https://www.aclweb.org/anthology/2021.naacl-main.43>

**AUTHORS:** Jonathan Herzig, Thomas M?ller, Syrine Krichene, Julian Eisenschlos

**HIGHLIGHT:** In this work, we tackle open-domain QA over tables for the first time, and show that retrieval can be improved by a retriever designed to handle tabular context.

44, **TITLE:** Open-Domain Question Answering Goes Conversational via Question Rewriting

<https://www.aclweb.org/anthology/2021.naacl-main.44>

**AUTHORS:** Raviteja Anantha, Svitlana Vakulenko, Zhucheng Tu, Shayne Longpre, Stephen Pulman, Srinivas Chappidi

**HIGHLIGHT:** We introduce a new dataset for Question Rewriting in Conversational Context (QRCC), which contains 14K conversations with 80K question-answer pairs.

45, **TITLE:** QA-GNN: Reasoning with Language Models and Knowledge Graphs for Question Answering

<https://www.aclweb.org/anthology/2021.naacl-main.45>

**AUTHORS:** Michihiro Yasunaga, Hongyu Ren, Antoine Bosselut, Percy Liang, Jure Leskovec

**HIGHLIGHT:** Here we propose a new model, QA-GNN, which addresses the above challenges through two key innovations: (i) relevance scoring, where we use LMs to estimate the importance of KG nodes relative to the given QA context, and (ii) joint reasoning, where we connect the QA context and KG to form a joint graph, and mutually update their representations through graph-based message passing.

46, **TITLE:** XOR QA: Cross-lingual Open-Retrieval Question Answering

<https://www.aclweb.org/anthology/2021.naacl-main.46>

**AUTHORS:** Akari Asai, Jungo Kasai, Jonathan Clark, Kenton Lee, Eunsol Choi, Hannaneh Hajishirzi

**HIGHLIGHT:** Based on this dataset, we introduce a task framework, called Cross-lingual Open-Retrieval Question Answering (XOR QA), that consists of three new tasks involving cross-lingual document retrieval from multilingual and English resources.

47, **TITLE:** SPARTA: Efficient Open-Domain Question Answering via Sparse Transformer Matching Retrieval

<https://www.aclweb.org/anthology/2021.naacl-main.47>

**AUTHORS:** Tiancheng Zhao, Xiaopeng Lu, Kyusong Lee

**HIGHLIGHT:** We introduce SPARTA, a novel neural retrieval method that shows great promise in performance, generalization, and interpretability for open-domain question answering.

48, TITLE: Implicitly Abusive Language - What does it actually look like and why are we not getting there?  
<https://www.aclweb.org/anthology/2021.naacl-main.48>  
AUTHORS: Michael Wiegand, Josef Ruppenhofer, Elisabeth Eder  
HIGHLIGHT: In this position paper, we explain why existing datasets make learning implicit abuse difficult and what needs to be changed in the design of such datasets.

49, TITLE: The Importance of Modeling Social Factors of Language: Theory and Practice  
<https://www.aclweb.org/anthology/2021.naacl-main.49>  
AUTHORS: Dirk Hovy, Diyi Yang  
HIGHLIGHT: In this position paper, we argue that the reason for the current limitations is a focus on information content while ignoring language's social factors.

50, TITLE: On learning and representing social meaning in NLP: a sociolinguistic perspective  
<https://www.aclweb.org/anthology/2021.naacl-main.50>  
AUTHORS: Dong Nguyen, Laura Rosseel, Jack Grieve  
HIGHLIGHT: We introduce the concept of social meaning to NLP and discuss how insights from sociolinguistics can inform work on representation learning in NLP.

51, TITLE: Preregistering NLP research  
<https://www.aclweb.org/anthology/2021.naacl-main.51>  
AUTHORS: Emiel van Miltenburg, Chris van der Lee, Emiel Kraemer  
HIGHLIGHT: The goal of this paper is to elicit a discussion in the NLP community, which we hope to synthesise into a general NLP preregistration form in future research.

52, TITLE: Get Your Vitamin C! Robust Fact Verification with Contrastive Evidence  
<https://www.aclweb.org/anthology/2021.naacl-main.52>  
AUTHORS: Tal Schuster, Adam Fisch, Regina Barzilay  
HIGHLIGHT: We present VitaminC, a benchmark infused with challenging cases that require fact verification models to discern and adjust to slight factual changes.

53, TITLE: Representing Numbers in NLP: a Survey and a Vision  
<https://www.aclweb.org/anthology/2021.naacl-main.53>  
AUTHORS: Avijit Thawani, Jay Pujara, Filip Ilievski, Pedro Szekely  
HIGHLIGHT: We synthesize best practices for representing numbers in text and articulate a vision for holistic numeracy in NLP, comprised of design trade-offs and a unified evaluation.

54, TITLE: Extending Multi-Document Summarization Evaluation to the Interactive Setting  
<https://www.aclweb.org/anthology/2021.naacl-main.54>  
AUTHORS: Ori Shapira, Ramakanth Pasunuru, Hadar Ronen, Mohit Bansal, Yael Amsterdamer, Ido Dagan  
HIGHLIGHT: In this paper, we develop an end-to-end evaluation framework for interactive summarization, focusing on expansion-based interaction, which considers the accumulating information along a user session.

55, TITLE: Identifying Helpful Sentences in Product Reviews  
<https://www.aclweb.org/anthology/2021.naacl-main.55>  
AUTHORS: Iftah Gamzu, Hila Gonen, Gilad Kutiel, Ran Levy, Eugene Agichtein  
HIGHLIGHT: In this work, we aim to maintain this advantage in situations where extreme brevity is needed, for example, when shopping by voice.

56, TITLE: Noisy Self-Knowledge Distillation for Text Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.56>  
AUTHORS: Yang Liu, Sheng Shen, Mirella Lapata  
HIGHLIGHT: In this paper we apply self-knowledge distillation to text summarization which we argue can alleviate problems with maximum-likelihood training on single reference and noisy datasets.

57, TITLE: Improving Zero and Few-Shot Abstractive Summarization with Intermediate Fine-tuning and Data Augmentation  
<https://www.aclweb.org/anthology/2021.naacl-main.57>  
AUTHORS: Alexander Fabbri, Simeng Han, Haoyuan Li, Haoran Li, Marjan Ghazvininejad, Shafiq Joty, Dragomir Radev, Yashar Mehdad

**HIGHLIGHT:** In this work, we introduce a novel and generalizable method, called WikiTransfer, for fine-tuning pretrained models for summarization in an unsupervised, dataset-specific manner.

58, **TITLE:** Enhancing Factual Consistency of Abstractive Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.58>  
**AUTHORS:** Chenguang Zhu, William Hinthorn, Ruochen Xu, Qingkai Zeng, Michael Zeng, Xuedong Huang, Meng Jiang  
**HIGHLIGHT:** We propose a fact-aware summarization model FASum to extract and integrate factual relations into the summary generation process via graph attention. We then design a factual corrector model FC to automatically correct factual errors from summaries generated by existing systems.

59, **TITLE:** Few-shot Intent Classification and Slot Filling with Retrieved Examples  
<https://www.aclweb.org/anthology/2021.naacl-main.59>  
**AUTHORS:** Dian Yu, Luheng He, Yuan Zhang, Xinya Du, Panupong Pasupat, Qi Li  
**HIGHLIGHT:** To this end, we propose a span-level retrieval method that learns similar contextualized representations for spans with the same label via a novel batch-softmax objective.

60, **TITLE:** "Nice Try, Kiddo": Investigating Ad Hominems in Dialogue Responses  
<https://www.aclweb.org/anthology/2021.naacl-main.60>  
**AUTHORS:** Emily Sheng, Kai-Wei Chang, Prem Natarajan, Nanyun Peng  
**HIGHLIGHT:** To this end, we propose categories of ad hominems, compose an annotated dataset, and build a classifier to analyze human and dialogue system responses to English Twitter posts.

61, **TITLE:** Human-like informative conversations: Better acknowledgements using conditional mutual information  
<https://www.aclweb.org/anthology/2021.naacl-main.61>  
**AUTHORS:** Ashwin Paranjape, Christopher Manning  
**HIGHLIGHT:** This work aims to build a dialogue agent that can weave new factual content into conversations as naturally as humans.

62, **TITLE:** A Comparative Study on Schema-Guided Dialogue State Tracking  
<https://www.aclweb.org/anthology/2021.naacl-main.62>  
**AUTHORS:** Jie Cao, Yi Zhang  
**HIGHLIGHT:** In this paper, we conduct in-depth comparative studies to understand the use of natural language description for schema in dialog state tracking.

63, **TITLE:** Spoken Language Understanding for Task-oriented Dialogue Systems with Augmented Memory Networks  
<https://www.aclweb.org/anthology/2021.naacl-main.63>  
**AUTHORS:** Jie Wu, Ian Harris, Hongzhi Zhao  
**HIGHLIGHT:** In this paper, we propose a novel approach to model long-term slot context and to fully utilize the semantic correlation between slots and intents.

64, **TITLE:** How to Motivate Your Dragon: Teaching Goal-Driven Agents to Speak and Act in Fantasy Worlds  
<https://www.aclweb.org/anthology/2021.naacl-main.64>  
**AUTHORS:** Prithviraj Ammanabrolu, Jack Urbanek, Margaret Li, Arthur Szlam, Tim Rocktschel, Jason Weston  
**HIGHLIGHT:** We introduce a reinforcement learning system that (1) incorporates large-scale language modeling-based and commonsense reasoning-based pre-training to imbue the agent with relevant priors; and (2) leverages a factorized action space of action commands and dialogue, balancing between the two.

65, **TITLE:** Linking Entities to Unseen Knowledge Bases with Arbitrary Schemas  
<https://www.aclweb.org/anthology/2021.naacl-main.65>  
**AUTHORS:** Yogarshi Vyas, Miguel Ballesteros  
**HIGHLIGHT:** This work focuses on linking to unseen KBs that do not have training data and whose schema is unknown during training.

66, **TITLE:** Self-Training with Weak Supervision  
<https://www.aclweb.org/anthology/2021.naacl-main.66>  
**AUTHORS:** Giannis Karamanolakis, Subhabrata Mukherjee, Guoqing Zheng, Ahmed Hassan Awadallah  
**HIGHLIGHT:** In this work, we develop a weak supervision framework (ASTRA) that leverages all the available data for a given task.

67, **TITLE:** Neural Language Modeling for Contextualized Temporal Graph Generation

<https://www.aclweb.org/anthology/2021.naacl-main.67>  
AUTHORS: Aman Madaan, Yiming Yang  
HIGHLIGHT: This paper presents the first study on using large-scale pre-trained language models for automated generation of an event-level temporal graph for a document.

68, TITLE: Probabilistic Box Embeddings for Uncertain Knowledge Graph Reasoning  
<https://www.aclweb.org/anthology/2021.naacl-main.68>  
AUTHORS: Xuelu Chen, Michael Boratko, Muhao Chen, Shib Sankar Dasgupta, Xiang Lorraine Li, Andrew McCallum  
HIGHLIGHT: To address these shortcomings, we propose BEUrRE, a novel uncertain knowledge graph embedding method with calibrated probabilistic semantics.

69, TITLE: Document-Level Event Argument Extraction by Conditional Generation  
<https://www.aclweb.org/anthology/2021.naacl-main.69>  
AUTHORS: Sha Li, Heng Ji, Jiawei Han  
HIGHLIGHT: We propose a document-level neural event argument extraction model by formulating the task as conditional generation following event templates.

70, TITLE: Template Filling with Generative Transformers  
<https://www.aclweb.org/anthology/2021.naacl-main.70>  
AUTHORS: Xinya Du, Alexander Rush, Claire Cardie  
HIGHLIGHT: We introduce a framework based on end-to-end generative transformers for this task (i.e., GTT).

71, TITLE: Towards Interpreting and Mitigating Shortcut Learning Behavior of NLU models  
<https://www.aclweb.org/anthology/2021.naacl-main.71>  
AUTHORS: Mengnan Du, Varun Manjunatha, Rajiv Jain, Ruchi Deshpande, Franck Démoncourt, Jiuxiang Gu, Tong Sun, Xia Hu  
HIGHLIGHT: In this work, we show that the words in the NLU training set can be modeled as a long-tailed distribution.

72, TITLE: On Attention Redundancy: A Comprehensive Study  
<https://www.aclweb.org/anthology/2021.naacl-main.72>  
AUTHORS: Yuchen Bian, Jiayi Huang, Xingyu Cai, Jiahong Yuan, Kenneth Church  
HIGHLIGHT: Based on the phase-independent and task-agnostic attention redundancy patterns, we propose a simple zero-shot pruning method as a case study.

73, TITLE: Does BERT Pretrained on Clinical Notes Reveal Sensitive Data?  
<https://www.aclweb.org/anthology/2021.naacl-main.73>  
AUTHORS: Eric Lehman, Sarthak Jain, Karl Pichotta, Yoav Goldberg, Byron Wallace  
HIGHLIGHT: In this work, we design a battery of approaches intended to recover Personal Health Information (PHI) from a trained BERT.

74, TITLE: Low-Complexity Probing via Finding Subnetworks  
<https://www.aclweb.org/anthology/2021.naacl-main.74>  
AUTHORS: Victor Sanh, Alexander Rush  
HIGHLIGHT: We instead propose a subtractive pruning-based probe, where we find an existing subnetwork that performs the linguistic task of interest.

75, TITLE: An Empirical Comparison of Instance Attribution Methods for NLP  
<https://www.aclweb.org/anthology/2021.naacl-main.75>  
AUTHORS: Pouya Pezeshkpour, Sarthak Jain, Byron Wallace, Sameer Singh  
HIGHLIGHT: In this work, we evaluate the degree to which different potential instance attribution agree with respect to the importance of training samples.

76, TITLE: Generalization in Instruction Following Systems  
<https://www.aclweb.org/anthology/2021.naacl-main.76>  
AUTHORS: Soham Dan, Michael Zhou, Dan Roth  
HIGHLIGHT: In this paper, we focus on instruction understanding in the blocks world domain and investigate the language understanding abilities of two top-performing systems for the task.

77, TITLE: LightningDOT: Pre-training Visual-Semantic Embeddings for Real-Time Image-Text Retrieval  
<https://www.aclweb.org/anthology/2021.naacl-main.77>



AUTHORS: Siqi Sun, Yen-Chun Chen, Linjie Li, Shuohang Wang, Yuwei Fang, Jingjing Liu  
HIGHLIGHT: We propose a simple yet highly effective approach, LightningDOT that accelerates the inference time of ITR by thousands of times, without sacrificing accuracy.

78, TITLE: Measuring Social Biases in Grounded Vision and Language Embeddings  
<https://www.aclweb.org/anthology/2021.naacl-main.78>  
AUTHORS: Candace Ross, Boris Katz, Andrei Barbu  
HIGHLIGHT: We introduce the space of generalizations (Grounded-WEAT and Grounded-SEAT) and demonstrate that three generalizations answer different yet important questions about how biases, language, and vision interact.

79, TITLE: MTAG: Modal-Temporal Attention Graph for Unaligned Human Multimodal Language Sequences  
<https://www.aclweb.org/anthology/2021.naacl-main.79>  
AUTHORS: Jianing Yang, Yongxin Wang, Ruitao Yi, Yuying Zhu, Azaan Rehman, Amir Zadeh, Soujanya Poria, Louis-Philippe Morency  
HIGHLIGHT: In this paper, we propose Modal-Temporal Attention Graph (MTAG).

80, TITLE: Grounding Open-Domain Instructions to Automate Web Support Tasks  
<https://www.aclweb.org/anthology/2021.naacl-main.80>  
AUTHORS: Nancy Xu, Sam Masling, Michael Du, Giovanni Campagna, Larry Heck, James Landay, Monica Lam  
HIGHLIGHT: We introduce a task and dataset to train AI agents from open-domain, step-by-step instructions originally written for people.

81, TITLE: Modular Networks for Compositional Instruction Following  
<https://www.aclweb.org/anthology/2021.naacl-main.81>  
AUTHORS: Rodolfo Corona, Daniel Fried, Coline Devin, Dan Klein, Trevor Darrell  
HIGHLIGHT: We propose a modular architecture for following natural language instructions that describe sequences of diverse subgoals.

82, TITLE: Improving Cross-Modal Alignment in Vision Language Navigation via Syntactic Information  
<https://www.aclweb.org/anthology/2021.naacl-main.82>  
AUTHORS: Jialu Li, Hao Tan, Mohit Bansal  
HIGHLIGHT: Hence, in this paper, we propose a navigation agent that utilizes syntax information derived from a dependency tree to enhance alignment between the instruction and the current visual scenes.

83, TITLE: Improving Pretrained Models for Zero-shot Multi-label Text Classification through Reinforced Label Hierarchy Reasoning  
<https://www.aclweb.org/anthology/2021.naacl-main.83>  
AUTHORS: Hui Liu, Danqing Zhang, Bing Yin, Xiaodan Zhu  
HIGHLIGHT: In this paper, we explore to improve pretrained models with label hierarchies on the ZS-MTC task.

84, TITLE: Fine-Tuning Pre-trained Language Model with Weak Supervision: A Contrastive-Regularized Self-Training Approach  
<https://www.aclweb.org/anthology/2021.naacl-main.84>  
AUTHORS: Yue Yu, Simiao Zuo, Haoming Jiang, Wendi Ren, Tuo Zhao, Chao Zhang  
HIGHLIGHT: We study the problem of fine-tuning pre-trained LMs using only weak supervision, without any labeled data.

85, TITLE: Posterior Differential Regularization with f-divergence for Improving Model Robustness  
<https://www.aclweb.org/anthology/2021.naacl-main.85>  
AUTHORS: Hao Cheng, Xiaodong Liu, Lis Pereira, Yaoliang Yu, Jianfeng Gao  
HIGHLIGHT: We address the problem of enhancing model robustness through regularization.

86, TITLE: Understanding Hard Negatives in Noise Contrastive Estimation  
<https://www.aclweb.org/anthology/2021.naacl-main.86>  
AUTHORS: Wenzheng Zhang, Karl Stratos  
HIGHLIGHT: We develop analytical tools to understand the role of hard negatives.

87, TITLE: Certified Robustness to Word Substitution Attack with Differential Privacy  
<https://www.aclweb.org/anthology/2021.naacl-main.87>  
AUTHORS: Wenjie Wang, Pengfei Tang, Jian Lou, Li Xiong

**HIGHLIGHT:** In this paper, we propose WordDP to achieve certified robustness against word substitution attacks in text classification via differential privacy (DP).

88, **TITLE:** DReCa: A General Task Augmentation Strategy for Few-Shot Natural Language Inference  
<https://www.aclweb.org/anthology/2021.naacl-main.88>

**AUTHORS:** Shikhar Murty, Tatsunori Hashimoto, Christopher Manning

**HIGHLIGHT:** To alleviate these issues, we propose DReCA (Decomposing datasets into Reasoning Categories), a simple method for discovering and using latent reasoning categories in a dataset, to form additional high quality tasks.

89, **TITLE:** Harnessing Multilinguality in Unsupervised Machine Translation for Rare Languages  
<https://www.aclweb.org/anthology/2021.naacl-main.89>

**AUTHORS:** Xavier Garcia, Aditya Siddhant, Orhan Firat, Ankur Parikh

**HIGHLIGHT:** In this work, we show that multilinguality is critical to making unsupervised systems practical for low-resource settings.

90, **TITLE:** Macro-Average: Rare Types Are Important Too  
<https://www.aclweb.org/anthology/2021.naacl-main.90>

**AUTHORS:** Thamme Gowda, Weiqiu You, Constantine Lignos, Jonathan May

**HIGHLIGHT:** We explore the simple type-based classifier metric, MacroF1, and study its applicability to MT evaluation.

91, **TITLE:** Assessing Reference-Free Peer Evaluation for Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.91>

**AUTHORS:** Sweta Agrawal, George Foster, Markus Freitag, Colin Cherry

**HIGHLIGHT:** We analyze various potential weaknesses of the approach, and find that it is surprisingly robust and likely to offer reasonable performance across a broad spectrum of domains and different system qualities.

92, **TITLE:** The Curious Case of Hallucinations in Neural Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.92>

**AUTHORS:** Vikas Raunak, Arul Menezes, Marcin Junczys-Dowmunt

**HIGHLIGHT:** In this work, we study hallucinations in Neural Machine Translation (NMT), which lie at an extreme end on the spectrum of NMT pathologies.

93, **TITLE:** Towards Continual Learning for Multilingual Machine Translation via Vocabulary Substitution  
<https://www.aclweb.org/anthology/2021.naacl-main.93>

**AUTHORS:** Xavier Garcia, Noah Constant, Ankur Parikh, Orhan Firat

**HIGHLIGHT:** We propose a straightforward vocabulary adaptation scheme to extend the language capacity of multilingual machine translation models, paving the way towards efficient continual learning for multilingual machine translation.

94, **TITLE:** Towards Modeling the Style of Translators in Neural Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.94>

**AUTHORS:** Yue Wang, Cuong Hoang, Marcello Federico

**HIGHLIGHT:** In this work, we investigate methods to augment the state of the art Transformer model with translator information that is available in part of the training data.

95, **TITLE:** Self-Supervised Test-Time Learning for Reading Comprehension  
<https://www.aclweb.org/anthology/2021.naacl-main.95>

**AUTHORS:** Pratyay Banerjee, Tejas Gokhale, Chitta Baral

**HIGHLIGHT:** In this work, we consider the task of unsupervised reading comprehension and present a method that performs "test-time learning" (TTL) on a given context (text passage), without requiring training on large-scale human-authored datasets containing context-question-answer triplets.

96, **TITLE:** Capturing Row and Column Semantics in Transformer Based Question Answering over Tables  
<https://www.aclweb.org/anthology/2021.naacl-main.96>

**AUTHORS:** Michael Glass, Mustafa Canim, Alfio Gliozzo, Saneem Chemmengath, Vishwajeet Kumar, Rishav Chakravarti, Avi Sil, Feifei Pan, Samarth Bharadwaj, Nicolas Rodolfo Fauceglia

**HIGHLIGHT:** In this paper, we propose two novel approaches demonstrating that one can achieve superior performance on table QA task without even using any of these specialized pre-training techniques.

97, **TITLE:** Explainable Multi-hop Verbal Reasoning Through Internal Monologue  
<https://www.aclweb.org/anthology/2021.naacl-main.97>

- AUTHORS: Zhengzhong Liang, Steven Bethard, Mihai Surdeanu  
HIGHLIGHT: We propose the Explainable multi-hop Verbal Reasoner (EVR) to solve these limitations by (a) decomposing multi-hop reasoning problems into several simple ones, and (b) using natural language to guide the intermediate reasoning hops.
- 98, TITLE: Robust Question Answering Through Sub-part Alignment  
<https://www.aclweb.org/anthology/2021.naacl-main.98>  
AUTHORS: Jifan Chen, Greg Durrett  
HIGHLIGHT: To make a more robust and understandable QA system, we model question answering as an alignment problem.
- 99, TITLE: Text Modular Networks: Learning to Decompose Tasks in the Language of Existing Models  
<https://www.aclweb.org/anthology/2021.naacl-main.99>  
AUTHORS: Tushar Khot, Daniel Khashabi, Kyle Richardson, Peter Clark, Ashish Sabharwal  
HIGHLIGHT: We propose a general framework called Text Modular Networks(TMNs) for building interpretable systems that learn to solve complex tasks by decomposing them into simpler ones solvable by existing models.
- 100, TITLE: RECONSIDER: Improved Re-Ranking using Span-Focused Cross-Attention for Open Domain Question Answering  
<https://www.aclweb.org/anthology/2021.naacl-main.100>  
AUTHORS: Srinivasan Iyer, Sewon Min, Yashar Mehdad, Wen-tau Yih  
HIGHLIGHT: We develop a successful re-ranking approach (RECONSIDER) for span-extraction tasks that improves upon the performance of MRC models, even beyond large-scale pre-training.
- 101, TITLE: On the Transferability of Minimal Prediction Preserving Inputs in Question Answering  
<https://www.aclweb.org/anthology/2021.naacl-main.101>  
AUTHORS: Shayne Longpre, Yi Lu, Chris DuBois  
HIGHLIGHT: In the context of question answering, we investigate competing hypotheses for the existence of MPPIs, including poor posterior calibration of neural models, lack of pretraining, and "dataset bias" (where a model learns to attend to spurious, non-generalizable cues in the training data).
- 102, TITLE: Understanding by Understanding Not: Modeling Negation in Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.102>  
AUTHORS: Arian Hosseini, Siva Reddy, Dzmitry Bahdanau, R Devon Hjelm, Alessandro Sordani, Aaron Courville  
HIGHLIGHT: To improve language models in this regard, we propose to augment the language modeling objective with an unlikelihood objective that is based on negated generic sentences from a raw text corpus.
- 103, TITLE: DuoRAT: Towards Simpler Text-to-SQL Models  
<https://www.aclweb.org/anthology/2021.naacl-main.103>  
AUTHORS: Torsten Scholak, Raymond Li, Dzmitry Bahdanau, Harm de Vries, Chris Pal  
HIGHLIGHT: Contrary to this trend, in this paper we focus on simplifications.
- 104, TITLE: Looking Beyond Sentence-Level Natural Language Inference for Question Answering and Text Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.104>  
AUTHORS: Anshuman Mishra, Dhruvesh Patel, Aparna Vijayakumar, Xiang Lorraine Li, Pavan Kapanipathi, Kartik Talamadupula  
HIGHLIGHT: In this work, we use the multiple-choice reading comprehension (MCRC) and checking factual correctness of textual summarization (CFCS) tasks to investigate potential reasons for this.
- 105, TITLE: Structure-Grounded Pretraining for Text-to-SQL  
<https://www.aclweb.org/anthology/2021.naacl-main.105>  
AUTHORS: Xiang Deng, Ahmed Hassan Awadallah, Christopher Meek, Oleksandr Polozov, Huan Sun, Matthew Richardson  
HIGHLIGHT: In this paper, we present a novel weakly supervised Structure-Grounded pretraining framework (STRUG) for text-to-SQL that can effectively learn to capture text-table alignment based on a parallel text-table corpus.
- 106, TITLE: Incremental Few-shot Text Classification with Multi-round New Classes: Formulation, Dataset and System  
<https://www.aclweb.org/anthology/2021.naacl-main.106>  
AUTHORS: Congying Xia, Wenpeng Yin, Yihao Feng, Philip Yu  
HIGHLIGHT: In this work, we define a new task in the NLP domain, incremental few-shot text classification, where the system incrementally handles multiple rounds of new classes.

- 107, TITLE: Temporal Reasoning on Implicit Events from Distant Supervision  
<https://www.aclweb.org/anthology/2021.naacl-main.107>  
AUTHORS: Ben Zhou, Kyle Richardson, Qiang Ning, Tushar Khot, Ashish Sabharwal, Dan Roth  
HIGHLIGHT: To address this, we propose a neuro-symbolic temporal reasoning model, SymTime, which exploits distant supervision signals from large-scale text and uses temporal rules to combine start times and durations to infer end times.
- 108, TITLE: Disentangling Semantics and Syntax in Sentence Embeddings with Pre-trained Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.108>  
AUTHORS: James Y. Huang, Kuan-Hao Huang, Kai-Wei Chang  
HIGHLIGHT: In this work, we present ParaBART, a semantic sentence embedding model that learns to disentangle semantics and syntax in sentence embeddings obtained by pre-trained language models.
- 109, TITLE: Structure-Aware Abstractive Conversation Summarization via Discourse and Action Graphs  
<https://www.aclweb.org/anthology/2021.naacl-main.109>  
AUTHORS: Jiaao Chen, Diyi Yang  
HIGHLIGHT: To this end, we propose to explicitly model the rich structures in conversations for more precise and accurate conversation summarization, by first incorporating discourse relations between utterances and action triples ("who-doing-what") in utterances through structured graphs to better encode conversations, and then designing a multi-granularity decoder to generate summaries by combining all levels of information.
- 110, TITLE: A New Approach to Overgenerating and Scoring Abstractive Summaries  
<https://www.aclweb.org/anthology/2021.naacl-main.110>  
AUTHORS: Kaiqiang Song, Bingqing Wang, Zhe Feng, Fei Liu  
HIGHLIGHT: We propose a new approach to generate multiple variants of the target summary with diverse content and varying lengths, then score and select admissible ones according to users' needs.
- 111, TITLE: D2S: Document-to-Slide Generation Via Query-Based Text Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.111>  
AUTHORS: Edward Sun, Yufang Hou, Dakuo Wang, Yunfeng Zhang, Nancy X. R. Wang  
HIGHLIGHT: In this work, we first contribute a new dataset, SciDuet, consisting of pairs of papers and their corresponding slides decks from recent years' NLP and ML conferences (e.g., ACL). Secondly, we present D2S, a novel system that tackles the document-to-slides task with a two-step approach:
- 112, TITLE: Efficient Attentions for Long Document Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.112>  
AUTHORS: Luyang Huang, Shuyang Cao, Nikolaus Parulian, Heng Ji, Lu Wang  
HIGHLIGHT: In this paper, we propose Hepos, a novel efficient encoder-decoder attention with head-wise positional strides to effectively pinpoint salient information from the source.
- 113, TITLE: RefSum: Refactoring Neural Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.113>  
AUTHORS: Yixin Liu, Zi-Yi Dou, Pengfei Liu  
HIGHLIGHT: In this work, we highlight several limitations of previous methods, which motivates us to present a new framework Refactor that provides a unified view of text summarization and summaries combination.
- 114, TITLE: Annotating and Modeling Fine-grained Factuality in Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.114>  
AUTHORS: Tanya Goyal, Greg Durrett  
HIGHLIGHT: We explore both synthetic and human-labeled data sources for training models to identify factual errors in summarization, and study factuality at the word-, dependency-, and sentence-level.
- 115, TITLE: Larger-Context Tagging: When and Why Does It Work?  
<https://www.aclweb.org/anthology/2021.naacl-main.115>  
AUTHORS: Jinlan Fu, Liangjing Feng, Qi Zhang, Xuanjing Huang, Pengfei Liu  
HIGHLIGHT: In this paper, instead of pursuing a state-of-the-art tagging system by architectural exploration, we focus on investigating when and why the larger-context training, as a general strategy, can work.
- 116, TITLE: Neural Sequence Segmentation as Determining the Leftmost Segments  
<https://www.aclweb.org/anthology/2021.naacl-main.116>  
AUTHORS: Yangming Li, Lema Liu, Kaisheng Yao

**HIGHLIGHT:** In this work, we propose a novel framework that incrementally segments natural language sentences at segment level.

117, **TITLE:** PCFGs Can Do Better: Inducing Probabilistic Context-Free Grammars with Many Symbols

<https://www.aclweb.org/anthology/2021.naacl-main.117>

**AUTHORS:** Songlin Yang, Yanpeng Zhao, Kewei Tu

**HIGHLIGHT:** In this work, we present a new parameterization form of PCFGs based on tensor decomposition, which has at most quadratic computational complexity in the symbol number and therefore allows us to use a much larger number of symbols.

118, **TITLE:** GEMNET: Effective Gated Gazetteer Representations for Recognizing Complex Entities in Low-context Input

<https://www.aclweb.org/anthology/2021.naacl-main.118>

**AUTHORS:** Tao Meng, Anjie Fang, Oleg Rokhlenko, Shervin Malmasi

**HIGHLIGHT:** We propose GEMNET, a novel approach for gazetteer knowledge integration, including (1) a flexible Contextual Gazetteer Representation (CGR) encoder that can be fused with any word-level model; and (2) a Mixture-of-Experts gating network that overcomes the feature overuse issue by learning to conditionally combine the context and gazetteer features, instead of assigning them fixed weights.

119, **TITLE:** Video-aided Unsupervised Grammar Induction

<https://www.aclweb.org/anthology/2021.naacl-main.119>

**AUTHORS:** Songyang Zhang, Linfeng Song, Lifeng Jin, Kun Xu, Dong Yu, Jiebo Luo

**HIGHLIGHT:** In this paper, we explore rich features (e.g. action, object, scene, audio, face, OCR and speech) from videos, taking the recent Compound PCFG model as the baseline.

120, **TITLE:** Generating Negative Samples by Manipulating Golden Responses for Unsupervised Learning of a Response Evaluation Model

<https://www.aclweb.org/anthology/2021.naacl-main.120>

**AUTHORS:** ChaeHun Park, Eugene Jang, Wonsuk Yang, Jong Park

**HIGHLIGHT:** For the unsupervised learning of such model, we propose a method of manipulating a golden response to create a new negative response that is designed to be inappropriate within the context while maintaining high similarity with the original golden response.

121, **TITLE:** How Robust are Fact Checking Systems on Colloquial Claims?

<https://www.aclweb.org/anthology/2021.naacl-main.121>

**AUTHORS:** Byeongchang Kim, Hyunwoo Kim, Seokhee Hong, Gunhee Kim

**HIGHLIGHT:** In this work, we ask: How robust are fact checking systems on claims in colloquial style?

122, **TITLE:** Fine-grained Post-training for Improving Retrieval-based Dialogue Systems

<https://www.aclweb.org/anthology/2021.naacl-main.122>

**AUTHORS:** Janghoon Han, Taesuk Hong, Byoungjae Kim, Youngjoong Ko, Jungyun Seo

**HIGHLIGHT:** To address this issue, we propose a new fine-grained post-training method that reflects the characteristics of the multi-turn dialogue.

123, **TITLE:** Put Chatbot into Its Interlocutor's Shoes: New Framework to Learn Chatbot Responding with Intention

<https://www.aclweb.org/anthology/2021.naacl-main.123>

**AUTHORS:** Hsuan Su, Jiun-Hao Jhan, Fan-yun Sun, Saurav Sahay, Hung-yi Lee

**HIGHLIGHT:** This paper proposes an innovative framework to train chatbots to possess human-like intentions.

124, **TITLE:** Adding Chit-Chat to Enhance Task-Oriented Dialogues

<https://www.aclweb.org/anthology/2021.naacl-main.124>

**AUTHORS:** Kai Sun, Seungwhan Moon, Paul Crook, Stephen Roller, Becka Silvert, Bing Liu, Zhiguang Wang, Honglei Liu, Eunjoon Cho, Claire Cardie

**HIGHLIGHT:** In this work, we propose to integrate both types of systems by Adding Chit-Chat to ENhance Task-ORiented dialogues (ACCENTOR), with the goal of making virtual assistant conversations more engaging and interactive.

125, **TITLE:** Incorporating Syntax and Semantics in Coreference Resolution with Heterogeneous Graph Attention Network

<https://www.aclweb.org/anthology/2021.naacl-main.125>

**AUTHORS:** Fan Jiang, Trevor Cohn

**HIGHLIGHT:** In this paper, we present a heterogeneous graph-based model to incorporate syntactic and semantic structures of sentences.

126, TITLE: Context Tracking Network: Graph-based Context Modeling for Implicit Discourse Relation Recognition  
<https://www.aclweb.org/anthology/2021.naacl-main.126>  
AUTHORS: Yingxue Zhang, Fandong Meng, Peng Li, Ping Jian, Jie Zhou  
HIGHLIGHT: In this paper, we thus propose a novel graph-based Context Tracking Network (CT-Net) to model the discourse context for IDRR.

127, TITLE: Improving Neural RST Parsing Model with Silver Agreement Subtrees  
<https://www.aclweb.org/anthology/2021.naacl-main.127>  
AUTHORS: Naoki Kobayashi, Tsutomu Hirao, Hidetaka Kamigaito, Manabu Okumura, Masaaki Nagata  
HIGHLIGHT: Therefore, we propose a method for improving neural RST parsing models by exploiting silver data, i.e., automatically annotated data.

128, TITLE: RST Parsing from Scratch  
<https://www.aclweb.org/anthology/2021.naacl-main.128>  
AUTHORS: Thanh-Tung Nguyen, Xuan-Phi Nguyen, Shafiq Joty, Xiaoli Li  
HIGHLIGHT: We introduce a novel top-down end-to-end formulation of document level discourse parsing in the Rhetorical Structure Theory (RST) framework.

129, TITLE: Did they answer? Subjective acts and intents in conversational discourse  
<https://www.aclweb.org/anthology/2021.naacl-main.129>  
AUTHORS: Elisa Ferracane, Greg Durrett, Junyi Jessy Li, Katrin Erk  
HIGHLIGHT: We present the first discourse dataset with multiple and subjective interpretations of English conversation in the form of perceived conversation acts and intents.

130, TITLE: Evaluating the Impact of a Hierarchical Discourse Representation on Entity Coreference Resolution Performance  
<https://www.aclweb.org/anthology/2021.naacl-main.130>  
AUTHORS: Sopan Khosla, James Fiacco, Carolyn Ros?  
HIGHLIGHT: In this work, we leverage automatically constructed discourse parse trees within a neural approach and demonstrate a significant improvement on two benchmark entity coreference-resolution datasets.

131, TITLE: Bridging Resolution: Making Sense of the State of the Art  
<https://www.aclweb.org/anthology/2021.naacl-main.131>  
AUTHORS: Hideo Kobayashi, Vincent Ng  
HIGHLIGHT: To shed light on these issues, we (1) propose a hybrid rule-based and MTL approach that would enable a better understanding of their comparative strengths and weaknesses; and (2) perform a manual analysis of the errors made by the MTL model.

132, TITLE: Explicitly Modeling Syntax in Language Models with Incremental Parsing and a Dynamic Oracle  
<https://www.aclweb.org/anthology/2021.naacl-main.132>  
AUTHORS: Yikang Shen, Shawn Tan, Alessandro Sordani, Siva Reddy, Aaron Courville  
HIGHLIGHT: In the present work, we propose a new syntax-aware language model: Syntactic Ordered Memory (SOM).

133, TITLE: Revisiting the Weaknesses of Reinforcement Learning for Neural Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.133>  
AUTHORS: Samuel Kiegele, Julia Kreutzer  
HIGHLIGHT: In this paper, we revisit these claims and study them under a wider range of configurations.

134, TITLE: Learning to Organize a Bag of Words into Sentences with Neural Networks: An Empirical Study  
<https://www.aclweb.org/anthology/2021.naacl-main.134>  
AUTHORS: Chongyang Tao, Shen Gao, Juntao Li, Yansong Feng, Dongyan Zhao, Rui Yan  
HIGHLIGHT: Inspired by such an intuition, in this paper, we perform a study to investigate how "order" information takes effects in natural language learning.

135, TITLE: Mask Attention Networks: Rethinking and Strengthen Transformer  
<https://www.aclweb.org/anthology/2021.naacl-main.135>  
AUTHORS: Zhihao Fan, Yeyun Gong, Dayiheng Liu, Zhongyu Wei, Siyuan Wang, Jian Jiao, Nan Duan, Ruofei Zhang, Xuanjing Huang  
HIGHLIGHT: In this paper, we present a novel understanding of SAN and FFN as Mask Attention Networks (MANs) and show that they are two special cases of MANs with static mask matrices.

- 136, TITLE: ERNIE-Gram: Pre-Training with Explicitly N-Gram Masked Language Modeling for Natural Language Understanding  
<https://www.aclweb.org/anthology/2021.naacl-main.136>  
AUTHORS: Dongling Xiao, Yu-Kun Li, Han Zhang, Yu Sun, Hao Tian, Hua Wu, Haifeng Wang  
HIGHLIGHT: As an alternative, we propose ERNIE-Gram, an explicitly n-gram masking method to enhance the integration of coarse-grained information into pre-training.
- 137, TITLE: Lattice-BERT: Leveraging Multi-Granularity Representations in Chinese Pre-trained Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.137>  
AUTHORS: Yuxuan Lai, Yijia Liu, Yansong Feng, Songfang Huang, Dongyan Zhao  
HIGHLIGHT: In this work, we propose a novel pre-training paradigm for Chinese - Lattice-BERT, which explicitly incorporates word representations along with characters, thus can model a sentence in a multi-granularity manner.
- 138, TITLE: Modeling Event Plausibility with Consistent Conceptual Abstraction  
<https://www.aclweb.org/anthology/2021.naacl-main.138>  
AUTHORS: Ian Porada, Kaheer Suleman, Adam Trischler, Jackie Chi Kit Cheung  
HIGHLIGHT: We find this inconsistency persists even when models are softly injected with lexical knowledge, and we present a simple post-hoc method of forcing model consistency that improves correlation with human plausibility judgements.
- 139, TITLE: UmlsBERT: Clinical Domain Knowledge Augmentation of Contextual Embeddings Using the Unified Medical Language System Metathesaurus  
<https://www.aclweb.org/anthology/2021.naacl-main.139>  
AUTHORS: George Michalopoulos, Yuanxin Wang, Hussam Kaka, Helen Chen, Alexander Wong  
HIGHLIGHT: We introduce UmlsBERT, a contextual embedding model that integrates domain knowledge during the pre-training process via a novel knowledge augmentation strategy.
- 140, TITLE: Field Embedding: A Unified Grain-Based Framework for Word Representation  
<https://www.aclweb.org/anthology/2021.naacl-main.140>  
AUTHORS: Junjie Luo, Xi Chen, Jichao Sun, Yuejia Xiang, Ningyu Zhang, Xiang Wan  
HIGHLIGHT: This work proposes a framework field embedding to jointly learn both word and grain embeddings by incorporating morphological, phonetic, and syntactical linguistic fields.
- 141, TITLE: MeBERT: Metaphor Detection via Contextualized Late Interaction using Metaphorical Identification Theories  
<https://www.aclweb.org/anthology/2021.naacl-main.141>  
AUTHORS: Minjin Choi, Sunkyung Lee, Eunseong Choi, Heesoo Park, Junhyuk Lee, Dongwon Lee, Jongwuk Lee  
HIGHLIGHT: To this end, we propose a novel metaphor detection model, namely metaphor-aware late interaction over BERT (MeBERT).
- 142, TITLE: Non-Parametric Few-Shot Learning for Word Sense Disambiguation  
<https://www.aclweb.org/anthology/2021.naacl-main.142>  
AUTHORS: Howard Chen, Mengzhou Xia, Danqi Chen  
HIGHLIGHT: In this work, we propose MetricWSD, a non-parametric few-shot learning approach to mitigate this data imbalance issue.
- 143, TITLE: Why Do Document-Level Polarity Classifiers Fail?  
<https://www.aclweb.org/anthology/2021.naacl-main.143>  
AUTHORS: Karen Martins, Pedro O.S Vaz-de-Melo, Rodrygo Santos  
HIGHLIGHT: We quantify the number of hard instances in polarity classification of movie reviews and provide empirical evidence about the need to pay attention to such problematic instances, as they are much harder to classify, for both machine and human classifiers.
- 144, TITLE: A Unified Span-Based Approach for Opinion Mining with Syntactic Constituents  
<https://www.aclweb.org/anthology/2021.naacl-main.144>  
AUTHORS: Qingrong Xia, Bo Zhang, Rui Wang, Zhenghua Li, Yue Zhang, Fei Huang, Luo Si, Min Zhang  
HIGHLIGHT: In this work, motivated by its span-based representations of opinion expressions and roles, we propose a unified span-based approach for the end-to-end OM setting.
- 145, TITLE: Target-specified Sequence Labeling with Multi-head Self-attention for Target-oriented Opinion Words Extraction  
<https://www.aclweb.org/anthology/2021.naacl-main.145>

AUTHORS: Yuhao Feng, Yanghui Rao, Yuyao Tang, Ninghua Wang, He Liu  
HIGHLIGHT: In this paper, we propose Target-Specified sequence labeling with Multi-head Self-Attention (TSMAS) for TOWE, in which any pre-trained language model with multi-head self-attention can be integrated conveniently.

146, TITLE: Does syntax matter? A strong baseline for Aspect-based Sentiment Analysis with RoBERTa  
<https://www.aclweb.org/anthology/2021.naacl-main.146>  
AUTHORS: Junqi Dai, Hang Yan, Tianxiang Sun, Pengfei Liu, Xipeng Qiu  
HIGHLIGHT: In this paper, we firstly compare the induced trees from PTMs and the dependency parsing trees on several popular models for the ABSA task, showing that the induced tree from fine-tuned RoBERTa (FT-RoBERTa) outperforms the parser-provided tree.

147, TITLE: Domain Divergences: A Survey and Empirical Analysis  
<https://www.aclweb.org/anthology/2021.naacl-main.147>  
AUTHORS: Abhinav Ramesh Kashyap, Devamanyu Hazarika, Min-Yen Kan, Roger Zimmermann  
HIGHLIGHT: We develop a taxonomy of divergence measures consisting of three classes - Information-theoretic, Geometric, and Higher-order measures and identify the relationships between them.

148, TITLE: Target-Aware Data Augmentation for Stance Detection  
<https://www.aclweb.org/anthology/2021.naacl-main.148>  
AUTHORS: Yingjie Li, Cornelia Caragea  
HIGHLIGHT: In this paper, we formulate the data augmentation of stance detection as a conditional masked language modeling task and augment the dataset by predicting the masked word conditioned on both its context and the auxiliary sentence that contains target and label information.

149, TITLE: End-to-end ASR to jointly predict transcriptions and linguistic annotations  
<https://www.aclweb.org/anthology/2021.naacl-main.149>  
AUTHORS: Motoi Omachi, Yuya Fujita, Shinji Watanabe, Matthew Wiesner  
HIGHLIGHT: We propose a Transformer-based sequence-to-sequence model for automatic speech recognition (ASR) capable of simultaneously transcribing and annotating audio with linguistic information such as phonemic transcripts or part-of-speech (POS) tags.

150, TITLE: Source and Target Bidirectional Knowledge Distillation for End-to-end Speech Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.150>  
AUTHORS: Hirofumi Inaguma, Tatsuya Kawahara, Shinji Watanabe  
HIGHLIGHT: In this work, we focus on sequence-level knowledge distillation (SeqKD) from external text-based NMT models.

151, TITLE: Searchable Hidden Intermediates for End-to-End Models of Decomposable Sequence Tasks  
<https://www.aclweb.org/anthology/2021.naacl-main.151>  
AUTHORS: Siddharth Dalmia, Brian Yan, Vikas Raunak, Florian Metze, Shinji Watanabe  
HIGHLIGHT: In this work, we present an end-to-end framework that exploits compositionality to learn searchable hidden representations at intermediate stages of a sequence model using decomposed sub-tasks.

152, TITLE: SPLAT: Speech-Language Joint Pre-Training for Spoken Language Understanding  
<https://www.aclweb.org/anthology/2021.naacl-main.152>  
AUTHORS: Yu-An Chung, Chenguang Zhu, Michael Zeng  
HIGHLIGHT: In this paper, we propose a novel semi-supervised learning framework, SPLAT, to jointly pre-train the speech and language modules.

153, TITLE: Worldly Wise (WoW) - Cross-Lingual Knowledge Fusion for Fact-based Visual Spoken-Question Answering  
<https://www.aclweb.org/anthology/2021.naacl-main.153>  
AUTHORS: Kiran Ramnath, Leda Sari, Mark Hasegawa-Johnson, Chang Yoo  
HIGHLIGHT: Towards these ends, we present a new task and a synthetically-generated dataset to do Fact-based Visual Spoken-Question Answering (FVSQA).

154, TITLE: Align-Refine: Non-Autoregressive Speech Recognition via Iterative Realignment  
<https://www.aclweb.org/anthology/2021.naacl-main.154>  
AUTHORS: Ethan A. Chi, Julian Salazar, Katrin Kirchhoff  
HIGHLIGHT: We instead propose \*iterative realignment\*, which by refining latent alignments allows more flexible edits in fewer steps.



- 155, TITLE: Everything Has a Cause: Leveraging Causal Inference in Legal Text Analysis  
<https://www.aclweb.org/anthology/2021.naacl-main.155>  
AUTHORS: Xiao Liu, Da Yin, Yansong Feng, Yuting Wu, Dongyan Zhao  
HIGHLIGHT: In this paper, we propose a novel Graph-based Causal Inference (GCI) framework, which builds causal graphs from fact descriptions without much human involvement and enables causal inference to facilitate legal practitioners to make proper decisions.
- 156, TITLE: Counterfactual Supporting Facts Extraction for Explainable Medical Record Based Diagnosis with Graph Network  
<https://www.aclweb.org/anthology/2021.naacl-main.156>  
AUTHORS: Haoran Wu, Wei Chen, Shuang Xu, Bo Xu  
HIGHLIGHT: Therefore, we propose a counterfactual multi-granularity graph supporting facts extraction (CMGE) method to extract supporting facts from irregular EMR itself without external knowledge bases in this paper.
- 157, TITLE: Personalized Response Generation via Generative Split Memory Network  
<https://www.aclweb.org/anthology/2021.naacl-main.157>  
AUTHORS: Yuwei Wu, Xuezhe Ma, Diyi Yang  
HIGHLIGHT: In this work, we look at how to generate personalized responses for questions on Reddit by utilizing personalized user profiles and posting histories.
- 158, TITLE: Towards Few-shot Fact-Checking via Perplexity  
<https://www.aclweb.org/anthology/2021.naacl-main.158>  
AUTHORS: Nayeon Lee, Yejin Bang, Andrea Madotto, Pascale Fung  
HIGHLIGHT: In this paper, we propose a new way of utilizing the powerful transfer learning ability of a language model via a perplexity score.
- 159, TITLE: Active2 Learning: Actively reducing redundancies in Active Learning methods for Sequence Tagging and Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.159>  
AUTHORS: Rishi Hazra, Parag Dutta, Shubham Gupta, Mohammed Abdul Qaathir, Ambedkar Dukkipati  
HIGHLIGHT: In this paper, we argue that since AL strategies choose examples independently, they may potentially select similar examples, all of which may not contribute significantly to the learning process.
- 160, TITLE: Generating An Optimal Interview Question Plan Using A Knowledge Graph And Integer Linear Programming  
<https://www.aclweb.org/anthology/2021.naacl-main.160>  
AUTHORS: Soham Datta, Prabir Mallick, Sangameshwar Patil, Indrajit Bhattacharya, Girish Palshikar  
HIGHLIGHT: We propose an interview assistant system to automatically, and in an objective manner, select an optimal set of technical questions (from question banks) personalized for a candidate.
- 161, TITLE: Model Extraction and Adversarial Transferability, Your BERT is Vulnerable!  
<https://www.aclweb.org/anthology/2021.naacl-main.161>  
AUTHORS: Xuanli He, Lingjuan Lyu, Lichao Sun, Qiongfai Xu  
HIGHLIGHT: In this work, we first present how an adversary can steal a BERT-based API service (the victim/target model) on multiple benchmark datasets with limited prior knowledge and queries. We further show that the extracted model can lead to highly transferable adversarial attacks against the victim model.
- 162, TITLE: A Global Past-Future Early Exit Method for Accelerating Inference of Pre-trained Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.162>  
AUTHORS: Kaiyuan Liao, Yi Zhang, Xuancheng Ren, Qi Su, Xu Sun, Bin He  
HIGHLIGHT: To address this issue, we propose a novel Past-Future method to make comprehensive predictions from a global perspective.
- 163, TITLE: Masked Conditional Random Fields for Sequence Labeling  
<https://www.aclweb.org/anthology/2021.naacl-main.163>  
AUTHORS: Tianwen Wei, Jianwei Qi, Shenghuan He, Songtao Sun  
HIGHLIGHT: In this work, we propose Masked Conditional Random Field (MCRF), an easy to implement variant of CRF that impose restrictions on candidate paths during both training and decoding phases.
- 164, TITLE: Heterogeneous Graph Neural Networks for Concept Prerequisite Relation Learning in Educational Data  
<https://www.aclweb.org/anthology/2021.naacl-main.164>

- AUTHORS: Chenghao Jia, Yongliang Shen, Yechun Tang, Lu Sun, Weiming Lu  
HIGHLIGHT: In this paper, we propose a novel concept prerequisite relation learning approach, named CPRL, which combines both concept representation learned from a heterogeneous graph and concept pairwise features.
- 165, TITLE: Be Careful about Poisoned Word Embeddings: Exploring the Vulnerability of the Embedding Layers in NLP Models  
<https://www.aclweb.org/anthology/2021.naacl-main.165>  
AUTHORS: Wenkai Yang, Lei Li, Zhiyuan Zhang, Xuancheng Ren, Xu Sun, Bin He  
HIGHLIGHT: However, in this paper, we find that it is possible to hack the model in a data-free way by modifying one single word embedding vector, with almost no accuracy sacrificed on clean samples.
- 166, TITLE: DA-Transformer: Distance-aware Transformer  
<https://www.aclweb.org/anthology/2021.naacl-main.166>  
AUTHORS: Chuhan Wu, Fangzhao Wu, Yongfeng Huang  
HIGHLIGHT: In this paper, we propose DA-Transformer, which is a distance-aware Transformer that can exploit the real distance.
- 167, TITLE: ASAP: A Chinese Review Dataset Towards Aspect Category Sentiment Analysis and Rating Prediction  
<https://www.aclweb.org/anthology/2021.naacl-main.167>  
AUTHORS: Jiahao Bu, Lei Ren, Shuang Zheng, Yang Yang, Jingang Wang, Fuzheng Zhang, Wei Wu  
HIGHLIGHT: To address the problem and advance related researches, we present a large-scale Chinese restaurant review dataset ASAP including 46, 730 genuine reviews from a leading online-to-offline (O2O) e-commerce platform in China.
- 168, TITLE: Are NLP Models really able to Solve Simple Math Word Problems?  
<https://www.aclweb.org/anthology/2021.naacl-main.168>  
AUTHORS: Arkil Patel, Satwik Bhattamishra, Navin Goyal  
HIGHLIGHT: In this paper, we restrict our attention to English MWPs taught in grades four and lower.
- 169, TITLE: WRIME: A New Dataset for Emotional Intensity Estimation with Subjective and Objective Annotations  
<https://www.aclweb.org/anthology/2021.naacl-main.169>  
AUTHORS: Tomoyuki Kajiwara, Chenhui Chu, Noriko Takemura, Yuta Nakashima, Hajime Nagahara  
HIGHLIGHT: In this study, we explore the difference between the emotional intensity of the writer and that of the readers with this dataset.
- 170, TITLE: KPQA: A Metric for Generative Question Answering Using Keyphrase Weights  
<https://www.aclweb.org/anthology/2021.naacl-main.170>  
AUTHORS: Hwanhee Lee, Seunghyun Yoon, Franck Dernoncourt, Doo Soon Kim, Trung Bui, Joongbo Shin, Kyomin Jung  
HIGHLIGHT: To alleviate this problem, we propose KPQA metric, a new metric for evaluating the correctness of GenQA.
- 171, TITLE: StylePTB: A Compositional Benchmark for Fine-grained Controllable Text Style Transfer  
<https://www.aclweb.org/anthology/2021.naacl-main.171>  
AUTHORS: Yiwei Lyu, Paul Pu Liang, Hai Pham, Eduard Hovy, Barnab? P?czos, Ruslan Salakhutdinov, Louis-Philippe Morency  
HIGHLIGHT: In this paper, we introduce a large-scale benchmark, StylePTB, with (1) paired sentences undergoing 21 fine-grained stylistic changes spanning atomic lexical, syntactic, semantic, and thematic transfers of text, as well as (2) compositions of multiple transfers which allow modeling of fine-grained stylistic changes as building blocks for more complex, high-level transfers.
- 172, TITLE: Blow the Dog Whistle: A Chinese Dataset for Cant Understanding with Common Sense and World Knowledge  
<https://www.aclweb.org/anthology/2021.naacl-main.172>  
AUTHORS: Canwen Xu, Wangchunshu Zhou, Tao Ge, Ke Xu, Julian McAuley, Furu Wei  
HIGHLIGHT: In this paper, we propose a large and diverse Chinese dataset for creating and understanding cant from a computational linguistics perspective.
- 173, TITLE: COVID-19 Named Entity Recognition for Vietnamese  
<https://www.aclweb.org/anthology/2021.naacl-main.173>  
AUTHORS: Thinh Hung Truong, Mai Hoang Dao, Dat Quoc Nguyen  
HIGHLIGHT: In this paper, we present the first manually-annotated COVID-19 domain-specific dataset for Vietnamese.
- 174, TITLE: Framing Unpacked: A Semi-Supervised Interpretable Multi-View Model of Media Frames  
<https://www.aclweb.org/anthology/2021.naacl-main.174>

**AUTHORS:** Shima Khanehzar, Trevor Cohn, Gosia Mikolajczak, Andrew Turpin, Lea Frermann  
**HIGHLIGHT:** This paper addresses both issues with a novel semi-supervised model, which jointly learns to embed local information about the events and related actors in a news article through an auto-encoding framework, and to leverage this signal for document-level frame classification.

**175, TITLE:** Automatic Classification of Neutralization Techniques in the Narrative of Climate Change Scepticism  
<https://www.aclweb.org/anthology/2021.naacl-main.175>  
**AUTHORS:** Shraey Bhatia, Jey Han Lau, Timothy Baldwin  
**HIGHLIGHT:** We first draw on social science to introduce the problem to the community of nlp, present the granularity of the coding schema and then collect manual annotations of neutralised techniques in text relating to climate change, and experiment with supervised and semi-supervised BERT-based models.

**176, TITLE:** Suicide Ideation Detection via Social and Temporal User Representations using Hyperbolic Learning  
<https://www.aclweb.org/anthology/2021.naacl-main.176>  
**AUTHORS:** Ramit Sawhney, Harshit Joshi, Rajiv Ratn Shah, Lucie Flek  
**HIGHLIGHT:** In this work, we propose a framework jointly leveraging a user's emotional history and social information from a user's neighborhood in a network to contextualize the interpretation of the latest tweet of a user on Twitter.

**177, TITLE:** WikiTalkEdit: A Dataset for modeling Editors' behaviors on Wikipedia  
<https://www.aclweb.org/anthology/2021.naacl-main.177>  
**AUTHORS:** Kokil Jaidka, Andrea Ceolin, Iknor Singh, Niyati Chhaya, Lyle Ungar  
**HIGHLIGHT:** This study introduces and analyzes WikiTalkEdit, a dataset of conversations and edit histories from Wikipedia, for research in online cooperation and conversation modeling.

**178, TITLE:** The structure of online social networks modulates the rate of lexical change  
<https://www.aclweb.org/anthology/2021.naacl-main.178>  
**AUTHORS:** Jian Zhu, David Jurgens  
**HIGHLIGHT:** We conduct a large-scale analysis of over 80k neologisms in 4420 online communities across a decade.

**179, TITLE:** Modeling Framing in Immigration Discourse on Social Media  
<https://www.aclweb.org/anthology/2021.naacl-main.179>  
**AUTHORS:** Julia Mendelsohn, Ceren Budak, David Jurgens  
**HIGHLIGHT:** By creating a new dataset of immigration-related tweets labeled for multiple framing typologies from political communication theory, we develop supervised models to detect frames.

**180, TITLE:** Modeling the Severity of Complaints in Social Media  
<https://www.aclweb.org/anthology/2021.naacl-main.180>  
**AUTHORS:** Mali Jin, Nikolaos Aletras  
**HIGHLIGHT:** In this paper, we study the severity level of complaints for the first time in computational linguistics.

**181, TITLE:** What About the Precedent: An Information-Theoretic Analysis of Common Law  
<https://www.aclweb.org/anthology/2021.naacl-main.181>  
**AUTHORS:** Josef Valvoda, Tiago Pimentel, Niklas Stoehr, Ryan Cotterell, Simone Teufel  
**HIGHLIGHT:** We are the first to approach this question computationally by comparing two longstanding jurisprudential views; Halsbury's, who believes that the arguments of the precedent are the main determinant of the outcome, and Goodhart's, who believes that what matters most is the precedent's facts.

**182, TITLE:** Introducing CAD: the Contextual Abuse Dataset  
<https://www.aclweb.org/anthology/2021.naacl-main.182>  
**AUTHORS:** Bertie Vidgen, Dong Nguyen, Helen Margetts, Patricia Rossini, Rebekah Tromble  
**HIGHLIGHT:** We introduce a new dataset of primarily English Reddit entries which addresses several limitations of prior work.

**183, TITLE:** Lifelong Learning of Hate Speech Classification on Social Media  
<https://www.aclweb.org/anthology/2021.naacl-main.183>  
**AUTHORS:** Jing Qian, Hong Wang, Mai ElSherief, Xifeng Yan  
**HIGHLIGHT:** In this work, we propose lifelong learning of hate speech classification on social media.

**184, TITLE:** Learning to Recognize Dialect Features  
<https://www.aclweb.org/anthology/2021.naacl-main.184>

AUTHORS: Dorottya Demszky, Devyani Sharma, Jonathan Clark, Vinodkumar Prabhakaran, Jacob Eisenstein  
HIGHLIGHT: In this paper, we introduce the task of dialect feature detection, and present two multitask learning approaches, both based on pretrained transformers.

185, TITLE: It's Not Just Size That Matters: Small Language Models Are Also Few-Shot Learners  
<https://www.aclweb.org/anthology/2021.naacl-main.185>  
AUTHORS: Timo Schick, Hinrich Sch?tze  
HIGHLIGHT: We show that performance similar to GPT-3 can be obtained with language models that are much "greener" in that their parameter count is several orders of magnitude smaller.

186, TITLE: Static Embeddings as Efficient Knowledge Bases?  
<https://www.aclweb.org/anthology/2021.naacl-main.186>  
AUTHORS: Philipp Duffer, Nora Kassner, Hinrich Sch?tze  
HIGHLIGHT: We show that, when restricting the output space to a candidate set, simple nearest neighbor matching using static embeddings performs better than PLMs.

187, TITLE: Highly Efficient Knowledge Graph Embedding Learning with Orthogonal Procrustes Analysis  
<https://www.aclweb.org/anthology/2021.naacl-main.187>  
AUTHORS: Xutan Peng, Guanyi Chen, Chenghua Lin, Mark Stevenson  
HIGHLIGHT: This paper proposes a simple yet effective KGE framework which can reduce the training time and carbon footprint by orders of magnitudes compared with state-of-the-art approaches, while producing competitive performance.

188, TITLE: Rethinking Network Pruning - under the Pre-train and Fine-tune Paradigm  
<https://www.aclweb.org/anthology/2021.naacl-main.188>  
AUTHORS: Dongkuan Xu, Ian En-Hsu Yen, Jinxi Zhao, Zhibin Xiao  
HIGHLIGHT: In this work, we aim to fill this gap by studying how knowledge are transferred and lost during the pre-train, fine-tune, and pruning process, and proposing a knowledge-aware sparse pruning process that achieves significantly superior results than existing literature.

189, TITLE: Towards a Comprehensive Understanding and Accurate Evaluation of Societal Biases in Pre-Trained Transformers  
<https://www.aclweb.org/anthology/2021.naacl-main.189>  
AUTHORS: Andrew Silva, Pradyumna Tambwekar, Matthew Gombolay  
HIGHLIGHT: In this paper, we investigate gender and racial bias across ubiquitous pre-trained language models, including GPT-2, XLNet, BERT, RoBERTa, ALBERT and DistilBERT.

190, TITLE: Detoxifying Language Models Risks Marginalizing Minority Voices  
<https://www.aclweb.org/anthology/2021.naacl-main.190>  
AUTHORS: Albert Xu, Eshaan Pathak, Eric Wallace, Suchin Gururangan, Maarten Sap, Dan Klein  
HIGHLIGHT: In this work, we show that these detoxification techniques hurt equity: they decrease the utility of LMs on language used by marginalized groups (e.g., African-American English and minority identity mentions).

191, TITLE: HONEST: Measuring Hurtful Sentence Completion in Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.191>  
AUTHORS: Debora Nozza, Federico Bianchi, Dirk Hovy  
HIGHLIGHT: We propose a score to measure hurtful sentence completions in language models (HONEST).

192, TITLE: EaSe: A Diagnostic Tool for VQA based on Answer Diversity  
<https://www.aclweb.org/anthology/2021.naacl-main.192>  
AUTHORS: Shailza Jolly, Sandro Pezzelle, Moin Nabi  
HIGHLIGHT: We propose EASE, a simple diagnostic tool for Visual Question Answering (VQA) which quantifies the difficulty of an image, question sample.

193, TITLE: DeCEMBERT: Learning from Noisy Instructional Videos via Dense Captions and Entropy Minimization  
<https://www.aclweb.org/anthology/2021.naacl-main.193>  
AUTHORS: Zineng Tang, Jie Lei, Mohit Bansal  
HIGHLIGHT: To address these issues, we propose an improved video-and-language pre-training method that first adds automatically-extracted dense region captions from the video frames as auxiliary text input, to provide informative visual cues for learning better video and language associations.

- 194, TITLE: Improving Generation and Evaluation of Visual Stories via Semantic Consistency  
<https://www.aclweb.org/anthology/2021.naacl-main.194>  
AUTHORS: Adyasha Maharana, Darryl Hannan, Mohit Bansal  
HIGHLIGHT: We present a number of improvements to prior modeling approaches, including (1) the addition of a dual learning framework that utilizes video captioning to reinforce the semantic alignment between the story and generated images, (2) a copy-transform mechanism for sequentially-consistent story visualization, and (3) MART-based transformers to model complex interactions between frames.
- 195, TITLE: Multilingual Multimodal Pre-training for Zero-Shot Cross-Lingual Transfer of Vision-Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.195>  
AUTHORS: Po-Yao Huang, Mandela Patrick, Junjie Hu, Graham Neubig, Florian Metze, Alexander Hauptmann  
HIGHLIGHT: Specifically, we focus on multilingual text-to-video search and propose a Transformer-based model that learns contextual multilingual multimodal embeddings.
- 196, TITLE: Video Question Answering with Phrases via Semantic Roles  
<https://www.aclweb.org/anthology/2021.naacl-main.196>  
AUTHORS: Arka Sadhu, Kan Chen, Ram Nevatia  
HIGHLIGHT: In this work, we leverage semantic roles derived from video descriptions to mask out certain phrases, to introduce VidQAP which poses VidQA as a fill-in-the-phrase task.
- 197, TITLE: From Masked Language Modeling to Translation: Non-English Auxiliary Tasks Improve Zero-shot Spoken Language Understanding  
<https://www.aclweb.org/anthology/2021.naacl-main.197>  
AUTHORS: Rob van der Goot, Ibrahim Sharaf, Aizhan Imankulova, Ahmet ?st?n, Marija Stepanovic, Alan Ramponi, Siti Oryza Khairunnisa, Mamoru Komachi, Barbara Plank  
HIGHLIGHT: To tackle the challenge, we propose a joint learning approach, with English SLU training data and non-English auxiliary tasks from raw text, syntax and translation for transfer.
- 198, TITLE: WEC: Deriving a Large-scale Cross-document Event Coreference dataset from Wikipedia  
<https://www.aclweb.org/anthology/2021.naacl-main.198>  
AUTHORS: Alon Eirew, Arie Cattan, Ido Dagan  
HIGHLIGHT: To complement these resources and enhance future research, we present Wikipedia Event Coreference (WEC), an efficient methodology for gathering a large-scale dataset for cross-document event coreference from Wikipedia, where coreference links are not restricted within predefined topics.
- 199, TITLE: Challenging distributional models with a conceptual network of philosophical terms  
<https://www.aclweb.org/anthology/2021.naacl-main.199>  
AUTHORS: Yvette Oortwijn, Jelke Bloem, Pia Sommerauer, Francois Meyer, Wei Zhou, Antske Fokkens  
HIGHLIGHT: We investigate the possibilities and limitations of using distributional semantic models for analyzing philosophical data by means of a realistic use-case.
- 200, TITLE: KILT: a Benchmark for Knowledge Intensive Language Tasks  
<https://www.aclweb.org/anthology/2021.naacl-main.200>  
AUTHORS: Fabio Petroni, Aleksandra Piktus, Angela Fan, Patrick Lewis, Majid Yazdani, Nicola De Cao, James Thorne, Yacine Jernite, Vladimir Karpukhin, Jean Maillard, Vassilis Plachouras, Tim Rockt?schel, Sebastian Riedel  
HIGHLIGHT: To catalyze research on models that condition on specific information in large textual resources, we present a benchmark for knowledge-intensive language tasks (KILT).
- 201, TITLE: A Survey on Recent Approaches for Natural Language Processing in Low-Resource Scenarios  
<https://www.aclweb.org/anthology/2021.naacl-main.201>  
AUTHORS: Michael A. Hedderich, Lukas Lange, Heike Adel, Jannik Str?tgen, Dietrich Klakow  
HIGHLIGHT: A goal of our survey is to explain how these methods differ in their requirements as understanding them is essential for choosing a technique suited for a specific low-resource setting.
- 202, TITLE: Temporal Knowledge Graph Completion using a Linear Temporal Regularizer and Multivector Embeddings  
<https://www.aclweb.org/anthology/2021.naacl-main.202>  
AUTHORS: Chengjin Xu, Yung-Yu Chen, Mojtaba Nayyeri, Jens Lehmann  
HIGHLIGHT: In this paper, we present a novel time-aware knowledge graph embedding approach, TeLM, which performs 4th-order tensor factorization of a Temporal knowledge graph using a Linear temporal regularizer and Multivector embeddings.
- 203, TITLE: UDALM: Unsupervised Domain Adaptation through Language Modeling

<https://www.aclweb.org/anthology/2021.naacl-main.203>

AUTHORS: Constantinos Karouzos, Georgios Paraskevopoulos, Alexandros Potamianos  
HIGHLIGHT: In this work we explore Unsupervised Domain Adaptation (UDA) of pretrained language models for downstream tasks.

204, TITLE: Beyond Black & White: Leveraging Annotator Disagreement via Soft-Label Multi-Task Learning

<https://www.aclweb.org/anthology/2021.naacl-main.204>

AUTHORS: Tommaso Fornaciari, Alexandra Uma, Silviu Paun, Barbara Plank, Dirk Hovy, Massimo Poesio  
HIGHLIGHT: We propose a novel method to incorporate this disagreement as information: in addition to the standard error computation, we use soft-labels (i.e., probability distributions over the annotator labels) as an auxiliary task in a multi-task neural network.

205, TITLE: Clustering-based Inference for Biomedical Entity Linking

<https://www.aclweb.org/anthology/2021.naacl-main.205>

AUTHORS: Rico Angell, Nicholas Monath, Sunil Mohan, Nishant Yadav, Andrew McCallum  
HIGHLIGHT: In this paper, we introduce a model in which linking decisions can be made not merely by linking to a knowledge base entity but also by grouping multiple mentions together via clustering and jointly making linking predictions.

206, TITLE: Variance-reduced First-order Meta-learning for Natural Language Processing Tasks

<https://www.aclweb.org/anthology/2021.naacl-main.206>

AUTHORS: Lingxiao Wang, Kevin Huang, Tengyu Ma, Quanquan Gu, Jing Huang  
HIGHLIGHT: Therefore, to address the overfitting issue when applying first-order meta-learning to NLP applications, we propose to reduce the variance of the gradient estimator used in task adaptation.

207, TITLE: Diversity-Aware Batch Active Learning for Dependency Parsing

<https://www.aclweb.org/anthology/2021.naacl-main.207>

AUTHORS: Tianze Shi, Adrian Benton, Igor Malioutov, Ozan Irsoy  
HIGHLIGHT: In this paper, we attempt to reduce the number of labeled examples needed to train a strong dependency parser using batch active learning (AL).

208, TITLE: How many data points is a prompt worth?

<https://www.aclweb.org/anthology/2021.naacl-main.208>

AUTHORS: Teven Le Scao, Alexander Rush  
HIGHLIGHT: We aim to quantify this benefit through rigorous testing of prompts in a fair setting: comparing prompted and head-based fine-tuning in equal conditions across many tasks and data sizes.

209, TITLE: Can Latent Alignments Improve Autoregressive Machine Translation?

<https://www.aclweb.org/anthology/2021.naacl-main.209>

AUTHORS: Adi Haviv, Lior Vassertail, Omer Levy  
HIGHLIGHT: We explore the possibility of training autoregressive machine translation models with latent alignment objectives, and observe that, in practice, this approach results in degenerate models.

210, TITLE: Smoothing and Shrinking the Sparse Seq2Seq Search Space

<https://www.aclweb.org/anthology/2021.naacl-main.210>

AUTHORS: Ben Peters, Andr  F. T. Martins  
HIGHLIGHT: In this work, we show that entmax-based models effectively solve the cat got your tongue problem, removing a major source of model error for neural machine translation.

211, TITLE: Unified Pre-training for Program Understanding and Generation

<https://www.aclweb.org/anthology/2021.naacl-main.211>

AUTHORS: Wasi Ahmad, Saikat Chakraborty, Baishakhi Ray, Kai-Wei Chang  
HIGHLIGHT: This paper introduces PLBART, a sequence-to-sequence model capable of performing a broad spectrum of program and language understanding and generation tasks.

212, TITLE: Hyperparameter-free Continuous Learning for Domain Classification in Natural Language Understanding

<https://www.aclweb.org/anthology/2021.naacl-main.212>

AUTHORS: Ting Hua, Yilin Shen, Changsheng Zhao, Yen-Chang Hsu, Hongxia Jin  
HIGHLIGHT: In this paper, we proposed a hyperparameter-free continual learning model for text data that can stably produce high performance under various environments.

- 213, TITLE: On the Embeddings of Variables in Recurrent Neural Networks for Source Code  
<https://www.aclweb.org/anthology/2021.naacl-main.213>  
AUTHORS: Nadezhda Chirkova  
HIGHLIGHT: In this work, we develop dynamic embeddings, a recurrent mechanism that adjusts the learned semantics of the variable when it obtains more information about the variable's role in the program.
- 214, TITLE: Cross-Lingual Word Embedding Refinement by  $l_1$  Norm Optimisation  
<https://www.aclweb.org/anthology/2021.naacl-main.214>  
AUTHORS: Xutan Peng, Chenghua Lin, Mark Stevenson  
HIGHLIGHT: Based on the more robust Manhattan norm (aka.  $l_1$  norm) goodness-of-fit criterion, this paper proposes a simple post-processing step to improve CLWEs.
- 215, TITLE: Semantic Frame Forecast  
<https://www.aclweb.org/anthology/2021.naacl-main.215>  
AUTHORS: Chieh-Yang Huang, Ting-Hao Huang  
HIGHLIGHT: This paper introduces Semantic Frame Forecast, a task that predicts the semantic frames that will occur in the next 10, 100, or even 1,000 sentences in a running story.
- 216, TITLE: MUSER: MULTImodal Stress detection using Emotion Recognition as an Auxiliary Task  
<https://www.aclweb.org/anthology/2021.naacl-main.216>  
AUTHORS: Yiqun Yao, Michalis Papakostas, Mihai Burzo, Mohamed Abouelenien, Rada Mihalcea  
HIGHLIGHT: In this work, we investigate the value of emotion recognition as an auxiliary task to improve stress detection.
- 217, TITLE: Learning to Decompose and Organize Complex Tasks  
<https://www.aclweb.org/anthology/2021.naacl-main.217>  
AUTHORS: Yi Zhang, Sujay Kumar Jauhar, Julia Kiseleva, Ryen White, Dan Roth  
HIGHLIGHT: Thus in this paper, we propose a novel end-to-end pipeline that consumes a complex task and induces a dependency graph from unstructured text to represent sub-tasks and their relationships.
- 218, TITLE: Continual Learning for Text Classification with Information Disentanglement Based Regularization  
<https://www.aclweb.org/anthology/2021.naacl-main.218>  
AUTHORS: Yufan Huang, Yanzhe Zhang, Jiaao Chen, Xuezhi Wang, Diyi Yang  
HIGHLIGHT: In this work, we propose an information disentanglement based regularization method for continual learning on text classification.
- 219, TITLE: Learning from Executions for Semantic Parsing  
<https://www.aclweb.org/anthology/2021.naacl-main.219>  
AUTHORS: Bailin Wang, Mirella Lapata, Ivan Titov  
HIGHLIGHT: In this work, we focus on the task of semi-supervised learning where a limited amount of annotated data is available together with many unlabeled NL utterances.
- 220, TITLE: Learning to Synthesize Data for Semantic Parsing  
<https://www.aclweb.org/anthology/2021.naacl-main.220>  
AUTHORS: Bailin Wang, Wenpeng Yin, Xi Victoria Lin, Caiming Xiong  
HIGHLIGHT: In this work, we propose a generative model which features a (non-neural) PCFG that models the composition of programs (e.g., SQL), and a BART-based translation model that maps a program to an utterance.
- 221, TITLE: Edge: Enriching Knowledge Graph Embeddings with External Text  
<https://www.aclweb.org/anthology/2021.naacl-main.221>  
AUTHORS: Saed Rezayi, Handong Zhao, Sungchul Kim, Ryan Rossi, Nedim Lipka, Sheng Li  
HIGHLIGHT: Previous work has partially addressed this issue by enriching knowledge graph entities based on "hard" co-occurrence of words present in the entities of the knowledge graphs and external text, while we achieve "soft" augmentation by proposing a knowledge graph enrichment and embedding framework named Edge.
- 222, TITLE: FLIN: A Flexible Natural Language Interface for Web Navigation  
<https://www.aclweb.org/anthology/2021.naacl-main.222>  
AUTHORS: Sahisnu Mazumder, Oriana Riva  
HIGHLIGHT: We propose FLIN, a natural language interface for web navigation that maps user commands to concept-level actions (rather than low-level UI actions), thus being able to flexibly adapt to different websites and handle their transient nature.

- 223, TITLE: Game-theoretic Vocabulary Selection via the Shapley Value and Banzhaf Index  
<https://www.aclweb.org/anthology/2021.naacl-main.223>  
AUTHORS: Roma Patel, Marta Garnelo, Ian Gemp, Chris Dyer, Yoram Bachrach  
HIGHLIGHT: We propose a vocabulary selection method that views words as members of a team trying to maximize the model's performance.
- 224, TITLE: Incorporating External Knowledge to Enhance Tabular Reasoning  
<https://www.aclweb.org/anthology/2021.naacl-main.224>  
AUTHORS: J. Neeraja, Vivek Gupta, Vivek Srikumar  
HIGHLIGHT: In this paper, we study these challenges through the problem of tabular natural language inference.
- 225, TITLE: Compositional Generalization for Neural Semantic Parsing via Span-level Supervised Attention  
<https://www.aclweb.org/anthology/2021.naacl-main.225>  
AUTHORS: Pengcheng Yin, Hao Fang, Graham Neubig, Adam Pauls, Emmanouil Antonios Platanios, Yu Su, Sam Thomson, Jacob Andreas  
HIGHLIGHT: We describe a span-level supervised attention loss that improves compositional generalization in semantic parsers.
- 226, TITLE: Domain Adaptation for Arabic Cross-Domain and Cross-Dialect Sentiment Analysis from Contextualized Word Embedding  
<https://www.aclweb.org/anthology/2021.naacl-main.226>  
AUTHORS: Abdellah El Mekki, Abdelkader El Mahdaouy, Ismail Berrada, Ahmed Khoumsi  
HIGHLIGHT: In this paper, we propose a new unsupervised domain adaptation method for Arabic cross-domain and cross-dialect sentiment analysis from Contextualized Word Embedding.
- 227, TITLE: Multi-task Learning of Negation and Speculation for Targeted Sentiment Classification  
<https://www.aclweb.org/anthology/2021.naacl-main.227>  
AUTHORS: Andrew Moore, Jeremy Barnes  
HIGHLIGHT: In this paper, we propose a multi-task learning method to incorporate information from syntactic and semantic auxiliary tasks, including negation and speculation scope detection, to create English-language models that are more robust to these phenomena.
- 228, TITLE: A Disentangled Adversarial Neural Topic Model for Separating Opinions from Plots in User Reviews  
<https://www.aclweb.org/anthology/2021.naacl-main.228>  
AUTHORS: Gabriele Pergola, Lin Gui, Yulan He  
HIGHLIGHT: In this paper, we propose a neural topic model combined with adversarial training to disentangle opinion topics from plot and neutral ones.
- 229, TITLE: Graph Ensemble Learning over Multiple Dependency Trees for Aspect-level Sentiment Classification  
<https://www.aclweb.org/anthology/2021.naacl-main.229>  
AUTHORS: Xiaochen Hou, Peng Qi, Guangtao Wang, Rex Ying, Jing Huang, Xiaodong He, Bowen Zhou  
HIGHLIGHT: To better leverage syntactic information in the face of unavoidable errors, we propose a simple yet effective graph ensemble technique, GraphMerge, to make use of the predictions from different parsers.
- 230, TITLE: Emotion-Infused Models for Explainable Psychological Stress Detection  
<https://www.aclweb.org/anthology/2021.naacl-main.230>  
AUTHORS: Elsbeth Turcan, Smaranda Muresan, Kathleen McKeown  
HIGHLIGHT: Here, we present work exploring the use of a semantically related task, emotion detection, for equally competent but more explainable and human-like psychological stress detection as compared to a black-box model.
- 231, TITLE: Aspect-based Sentiment Analysis with Type-aware Graph Convolutional Networks and Layer Ensemble  
<https://www.aclweb.org/anthology/2021.naacl-main.231>  
AUTHORS: Yuanhe Tian, Guimin Chen, Yan Song  
HIGHLIGHT: To address such limitations, in this paper, we propose an approach to explicitly utilize dependency types for ABSA with type-aware graph convolutional networks (T-GCN), where attention is used in T-GCN to distinguish different edges (relations) in the graph and attentive layer ensemble is proposed to comprehensively learn from different layers of T-GCN.
- 232, TITLE: Supertagging-based Parsing with Linear Context-free Rewriting Systems  
<https://www.aclweb.org/anthology/2021.naacl-main.232>  
AUTHORS: Thomas Ruprecht, Richard M?rbitz  
HIGHLIGHT: We present the first supertagging-based parser for linear context-free rewriting systems (LCFRS).



- 233, TITLE: Outside Computation with Superior Functions  
<https://www.aclweb.org/anthology/2021.naacl-main.233>  
AUTHORS: Parker Riley, Daniel Gildea  
HIGHLIGHT: We show that a general algorithm for efficient computation of outside values under the minimum of superior functions framework proposed by Knuth (1977) would yield a sub-exponential time algorithm for SAT, violating the Strong Exponential Time Hypothesis (SETH).
- 234, TITLE: Learning Syntax from Naturally-Occurring Bracketings  
<https://www.aclweb.org/anthology/2021.naacl-main.234>  
AUTHORS: Tianze Shi, Ozan Irsoy, Igor Malioutov, Lillian Lee  
HIGHLIGHT: But they are noisy and incomplete; to address this challenge, we develop a partial-brackets-aware structured ramp loss in learning.
- 235, TITLE: Bot-Adversarial Dialogue for Safe Conversational Agents  
<https://www.aclweb.org/anthology/2021.naacl-main.235>  
AUTHORS: Jing Xu, Da Ju, Margaret Li, Y-Lan Boureau, Jason Weston, Emily Dinan  
HIGHLIGHT: We introduce a new human-and-model-in-the-loop framework for evaluating the toxicity of such models, and compare a variety of existing methods in both the cases of non-adversarial and adversarial users that expose their weaknesses.
- 236, TITLE: Non-Autoregressive Semantic Parsing for Compositional Task-Oriented Dialog  
<https://www.aclweb.org/anthology/2021.naacl-main.236>  
AUTHORS: Arun Babu, Akshat Shrivastava, Armen Aghajanyan, Ahmed Aly, Angela Fan, Marjan Ghazvininejad  
HIGHLIGHT: In this work, we propose a non-autoregressive approach to predict semantic parse trees with an efficient seq2seq model architecture.
- 237, TITLE: Example-Driven Intent Prediction with Observers  
<https://www.aclweb.org/anthology/2021.naacl-main.237>  
AUTHORS: Shikib Mehri, Mihail Eric  
HIGHLIGHT: In this paper, we focus on the intent classification problem which aims to identify user intents given utterances addressed to the dialog system.
- 238, TITLE: Imperfect also Deserves Reward: Multi-Level and Sequential Reward Modeling for Better Dialog Management  
<https://www.aclweb.org/anthology/2021.naacl-main.238>  
AUTHORS: Zhengxu Hou, Bang Liu, Ruihui Zhao, Zijing Ou, Yafei Liu, Xi Chen, Yefeng Zheng  
HIGHLIGHT: In this paper, we propose a multi-level reward modeling approach that factorizes a reward into a three-level hierarchy: domain, act, and slot.
- 239, TITLE: Action-Based Conversations Dataset: A Corpus for Building More In-Depth Task-Oriented Dialogue Systems  
<https://www.aclweb.org/anthology/2021.naacl-main.239>  
AUTHORS: Derek Chen, Howard Chen, Yi Yang, Alexander Lin, Zhou Yu  
HIGHLIGHT: To study customer service dialogue systems in more realistic settings, we introduce the Action-Based Conversations Dataset (ABCD), a fully-labeled dataset with over 10K human-to-human dialogues containing 55 distinct user intents requiring unique sequences of actions constrained by policies to achieve task success.
- 240, TITLE: Controlling Dialogue Generation with Semantic Exemplars  
<https://www.aclweb.org/anthology/2021.naacl-main.240>  
AUTHORS: Prakhar Gupta, Jeffrey Bigham, Yulia Tsvetkov, Amy Pavel  
HIGHLIGHT: We present an Exemplar-based Dialogue Generation model, EDGE, that uses the semantic frames present in exemplar responses to guide response generation.
- 241, TITLE: COIL: Revisit Exact Lexical Match in Information Retrieval with Contextualized Inverted List  
<https://www.aclweb.org/anthology/2021.naacl-main.241>  
AUTHORS: Luyu Gao, Zhuyun Dai, Jamie Callan  
HIGHLIGHT: This paper presents COIL, a contextualized exact match retrieval architecture, where scoring is based on overlapping query document tokens' contextualized representations.
- 242, TITLE: X-Class: Text Classification with Extremely Weak Supervision  
<https://www.aclweb.org/anthology/2021.naacl-main.242>  
AUTHORS: Zihan Wang, Dheeraj Mekala, Jingbo Shang

**HIGHLIGHT:** In this paper, we explore text classification with extremely weak supervision, i.e., only relying on the surface text of class names.

243, **TITLE:** Fine-tuning Encoders for Improved Monolingual and Zero-shot Polylingual Neural Topic Modeling  
<https://www.aclweb.org/anthology/2021.naacl-main.243>

**AUTHORS:** Aaron Mueller, Mark Dredze

**HIGHLIGHT:** Thus, we propose several methods for fine-tuning encoders to improve both monolingual and zero-shot polylingual neural topic modeling.

244, **TITLE:** Exploring the Relationship Between Algorithm Performance, Vocabulary, and Run-Time in Text Classification  
<https://www.aclweb.org/anthology/2021.naacl-main.244>

**AUTHORS:** Wilson Fearn, Orion Weller, Kevin Seppi

**HIGHLIGHT:** To fill this gap, we provide a comprehensive study that examines how preprocessing techniques affect the vocabulary size, model performance, and model run-time, evaluating ten techniques over four models and two datasets.

245, **TITLE:** Faithfully Explainable Recommendation via Neural Logic Reasoning

<https://www.aclweb.org/anthology/2021.naacl-main.245>

**AUTHORS:** Yaxin Zhu, Yikun Xian, Zuohui Fu, Gerard de Melo, Yongfeng Zhang

**HIGHLIGHT:** Specifically, we propose neural logic reasoning for explainable recommendation (LOGGER) by drawing on interpretable logical rules to guide the path-reasoning process for explanation generation.

246, **TITLE:** You Sound Like Someone Who Watches Drama Movies: Towards Predicting Movie Preferences from Conversational Interactions

<https://www.aclweb.org/anthology/2021.naacl-main.246>

**AUTHORS:** Sergey Volokhin, Joyce Ho, Oleg Rokhlenko, Eugene Agichtein

**HIGHLIGHT:** We explore one promising direction for conversational recommendation: mapping a conversational user, for whom there is limited or no data available, to most similar external reviewers, whose preferences are known, by representing the conversation as a user's interest vector, and adapting collaborative filtering techniques to estimate the current user's preferences for new movies.

247, **TITLE:** Reading and Acting while Blindfolded: The Need for Semantics in Text Game Agents

<https://www.aclweb.org/anthology/2021.naacl-main.247>

**AUTHORS:** Shunyu Yao, Karthik Narasimhan, Matthew Hausknecht

**HIGHLIGHT:** To remedy this deficiency, we propose an inverse dynamics decoder to regularize the representation space and encourage exploration, which shows improved performance on several games including Zork I.

248, **TITLE:** SoRT-ing VQA Models : Contrastive Gradient Learning for Improved Consistency

<https://www.aclweb.org/anthology/2021.naacl-main.248>

**AUTHORS:** Sameer Dharur, Purva Tendulkar, Dhruv Batra, Devi Parikh, Ramprasaath R. Selvaraju

**HIGHLIGHT:** To address this, we first present a gradient-based interpretability approach to determine the questions most strongly correlated with the reasoning question on an image, and use this to evaluate VQA models on their ability to identify the relevant sub-questions needed to answer a reasoning question.

249, **TITLE:** Semi-Supervised Policy Initialization for Playing Games with Language Hints

<https://www.aclweb.org/anthology/2021.naacl-main.249>

**AUTHORS:** Tsu-Jui Fu, William Yang Wang

**HIGHLIGHT:** In this paper, we propose semi-supervised initialization (SSI) that allows the agent to learn from various possible hints before training under different tasks.

250, **TITLE:** Revisiting Document Representations for Large-Scale Zero-Shot Learning

<https://www.aclweb.org/anthology/2021.naacl-main.250>

**AUTHORS:** Jihyung Kil, Wei-Lun Chao

**HIGHLIGHT:** In this paper, we revisit the use of documents as semantic representations.

251, **TITLE:** Negative language transfer in learner English: A new dataset

<https://www.aclweb.org/anthology/2021.naacl-main.251>

**AUTHORS:** Leticia Farias Wanderley, Nicole Zhao, Carrie Demmans Epp

**HIGHLIGHT:** This paper introduces a learner English dataset in which learner errors are accompanied by information about possible error sources.

- 252, TITLE: SentSim: Crosslingual Semantic Evaluation of Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.252>  
AUTHORS: Yurun Song, Junchen Zhao, Lucia Specia  
HIGHLIGHT: In this paper, we propose a more cost-effective, yet well performing unsupervised alternative SentSim: relying on strong pretrained multilingual word and sentence representations, we directly compare the source with the machine translated sentence, thus avoiding the need for both reference translations and labelled training data.
- 253, TITLE: Quality Estimation for Image Captions Based on Large-scale Human Evaluations  
<https://www.aclweb.org/anthology/2021.naacl-main.253>  
AUTHORS: Tomer Levinboim, Ashish V. Thapliyal, Piyush Sharma, Radu Soricut  
HIGHLIGHT: In this paper, we focus on the task of Quality Estimation (QE) for image captions, which attempts to model the caption quality from a human perspective and \*without\* access to ground-truth references, so that it can be applied at prediction time to detect low-quality captions produced on \*previously unseen images\*.
- 254, TITLE: CaSiNo: A Corpus of Campsite Negotiation Dialogues for Automatic Negotiation Systems  
<https://www.aclweb.org/anthology/2021.naacl-main.254>  
AUTHORS: Kushal Chawla, Jaysa Ramirez, Rene Clever, Gale Lucas, Jonathan May, Jonathan Gratch  
HIGHLIGHT: To advance the development of practical negotiation systems, we present CaSiNo: a novel corpus of over a thousand negotiation dialogues in English.
- 255, TITLE: News Headline Grouping as a Challenging NLU Task  
<https://www.aclweb.org/anthology/2021.naacl-main.255>  
AUTHORS: Philippe Laban, Lucas Bandarkar, Marti A. Hearst  
HIGHLIGHT: In this paper, we introduce the task of HeadLine Grouping (HLG) and a corresponding dataset (HLGD) consisting of 20,056 pairs of news headlines, each labeled with a binary judgement as to whether the pair belongs within the same group.
- 256, TITLE: O!?, Bonjour, Salve! XFORMAL: A Benchmark for Multilingual Formality Style Transfer  
<https://www.aclweb.org/anthology/2021.naacl-main.256>  
AUTHORS: Eleftheria Briakou, Di Lu, Ke Zhang, Joel Tetreault  
HIGHLIGHT: We take the first step towards multilingual style transfer by creating and releasing XFORMAL, a benchmark of multiple formal reformulations of informal text in Brazilian Portuguese, French, and Italian.
- 257, TITLE: Grouping Words with Semantic Diversity  
<https://www.aclweb.org/anthology/2021.naacl-main.257>  
AUTHORS: Karine Chubarian, Abdul Rafae Khan, Anastasios Sidiropoulos, Jia Xu  
HIGHLIGHT: We introduce an approach by grouping input words based on their semantic diversity to simplify input language representation with low ambiguity.
- 258, TITLE: Noise Stability Regularization for Improving BERT Fine-tuning  
<https://www.aclweb.org/anthology/2021.naacl-main.258>  
AUTHORS: Hang Hua, Xingjian Li, Dejing Dou, Chengzhong Xu, Jiebo Luo  
HIGHLIGHT: Specifically, we introduce a novel and effective regularization method to improve fine-tuning on NLP tasks, referred to as Layer-wise Noise Stability Regularization (LNSR).
- 259, TITLE: FlowPrior: Learning Expressive Priors for Latent Variable Sentence Models  
<https://www.aclweb.org/anthology/2021.naacl-main.259>  
AUTHORS: Xiaoran Ding, Kevin Gimpel  
HIGHLIGHT: We find that existing training strategies are not effective for learning rich priors, so we propose adding the importance-sampled log marginal likelihood as a second term to the standard VAE objective to help when learning the prior.
- 260, TITLE: HTCInfoMax: A Global Model for Hierarchical Text Classification via Information Maximization  
<https://www.aclweb.org/anthology/2021.naacl-main.260>  
AUTHORS: Zhongfen Deng, Hao Peng, Dongxiao He, Jianxin Li, Philip Yu  
HIGHLIGHT: In this paper, we propose HTCInfoMax to address these issues by introducing information maximization which includes two modules: text-label mutual information maximization and label prior matching.
- 261, TITLE: Knowledge Guided Metric Learning for Few-Shot Text Classification  
<https://www.aclweb.org/anthology/2021.naacl-main.261>  
AUTHORS: Dianbo Sui, Yubo Chen, Binjie Mao, Delai Qiu, Kang Liu, Jun Zhao

- HIGHLIGHT:** Inspired by human intelligence, we propose to introduce external knowledge into few-shot learning to imitate human knowledge.
- 262, **TITLE:** Ensemble of MRR and NDCG models for Visual Dialog  
<https://www.aclweb.org/anthology/2021.naacl-main.262>  
**AUTHORS:** Idan Schwartz  
**HIGHLIGHT:** To address this issue, we describe a two-step non-parametric ranking approach that can merge strong MRR and NDCG models.
- 263, **TITLE:** Supervised Neural Clustering via Latent Structured Output Learning: Application to Question Intents  
<https://www.aclweb.org/anthology/2021.naacl-main.263>  
**AUTHORS:** Iryna Haponchyk, Alessandro Moschitti  
**HIGHLIGHT:** In this paper, we design neural networks based on latent structured prediction loss and Transformer models to approach supervised clustering.
- 264, **TITLE:** ConVEx: Data-Efficient and Few-Shot Slot Labeling  
<https://www.aclweb.org/anthology/2021.naacl-main.264>  
**AUTHORS:** Mathew Henderson, Ivan Vulic  
**HIGHLIGHT:** We propose ConVEx (Conversational Value Extractor), an efficient pretraining and fine-tuning neural approach for slot-labeling dialog tasks.
- 265, **TITLE:** CREAD: Combined Resolution of Ellipses and Anaphora in Dialogues  
<https://www.aclweb.org/anthology/2021.naacl-main.265>  
**AUTHORS:** Bo-Hsiang Tseng, Shruti Bhargava, Jiarui Lu, Joel Ruben Antony Moniz, Dhivya Piraviperumal, Lin Li, Hong Yu  
**HIGHLIGHT:** In this work, we propose a novel joint learning framework of modeling coreference resolution and query rewriting for complex, multi-turn dialogue understanding.
- 266, **TITLE:** Knowledge-Driven Slot Constraints for Goal-Oriented Dialogue Systems  
<https://www.aclweb.org/anthology/2021.naacl-main.266>  
**AUTHORS:** Piyawat Lertvittayakumjorn, Daniele Bonadiman, Saab Mansour  
**HIGHLIGHT:** In this paper, we formalize knowledge-driven slot constraints and present a new task of constraint violation detection accompanied with benchmarking data.
- 267, **TITLE:** Clipping Loops for Sample-Efficient Dialogue Policy Optimisation  
<https://www.aclweb.org/anthology/2021.naacl-main.267>  
**AUTHORS:** Yen-Chen Wu, Carl Edward Rasmussen  
**HIGHLIGHT:** In this paper, we propose loop-clipping policy optimisation (LCPO) to eliminate useless responses.
- 268, **TITLE:** Integrating Lexical Information into Entity Neighbourhood Representations for Relation Prediction  
<https://www.aclweb.org/anthology/2021.naacl-main.268>  
**AUTHORS:** Ian Wood, Mark Johnson, Stephen Wan  
**HIGHLIGHT:** We present an extension of OpenKi that incorporates embeddings of text-based representations of the entities and the relations.
- 269, **TITLE:** Noisy-Labeled NER with Confidence Estimation  
<https://www.aclweb.org/anthology/2021.naacl-main.269>  
**AUTHORS:** Kun Liu, Yao Fu, Chuanqi Tan, Mosha Chen, Ningyu Zhang, Songfang Huang, Sheng Gao  
**HIGHLIGHT:** Based on empirical observations of different training dynamics of noisy and clean labels, we propose strategies for estimating confidence scores based on local and global independence assumptions.
- 270, **TITLE:** TABBIE: Pretrained Representations of Tabular Data  
<https://www.aclweb.org/anthology/2021.naacl-main.270>  
**AUTHORS:** Hiroshi Iida, Dung Thai, Varun Manjunatha, Mohit Iyyer  
**HIGHLIGHT:** We devise a simple pretraining objective (corrupt cell detection) that learns exclusively from tabular data and reaches the state-of-the-art on a suite of table-based prediction tasks.
- 271, **TITLE:** Better Feature Integration for Named Entity Recognition  
<https://www.aclweb.org/anthology/2021.naacl-main.271>  
**AUTHORS:** Lu Xu, Zhanming Jie, Wei Lu, Lidong Bing

**HIGHLIGHT:** In this work, we propose a simple and robust solution to incorporate both types of features with our Synergized-LSTM (Syn-LSTM), which clearly captures how the two types of features interact.

272, **TITLE:** ZS-BERT: Towards Zero-Shot Relation Extraction with Attribute Representation Learning

<https://www.aclweb.org/anthology/2021.naacl-main.272>

**AUTHORS:** Chih-Yao Chen, Cheng-Te Li

**HIGHLIGHT:** In this paper, we formulate the zero-shot relation extraction problem by incorporating the text description of seen and unseen relations.

273, **TITLE:** Graph Convolutional Networks for Event Causality Identification with Rich Document-level Structures

<https://www.aclweb.org/anthology/2021.naacl-main.273>

**AUTHORS:** Minh Tran Phu, Thien Huu Nguyen

**HIGHLIGHT:** As such, we propose a graph-based model that constructs interaction graphs to capture relevant connections between important objects for DECI in input documents.

274, **TITLE:** A Context-Dependent Gated Module for Incorporating Symbolic Semantics into Event Coreference Resolution

<https://www.aclweb.org/anthology/2021.naacl-main.274>

**AUTHORS:** Tuan Lai, Heng Ji, Trung Bui, Quan Hung Tran, Franck Demoncourt, Walter Chang

**HIGHLIGHT:** Motivated by these observations, we propose a novel context-dependent gated module to adaptively control the information flows from the input symbolic features.

275, **TITLE:** Multi-Style Transfer with Discriminative Feedback on Disjoint Corpus

<https://www.aclweb.org/anthology/2021.naacl-main.275>

**AUTHORS:** Navita Goyal, Balaji Vasan Srinivasan, Anandhavelu N, Abhilasha Sancheti

**HIGHLIGHT:** In our work, we relax this requirement of jointly annotated data across multiple styles by using independently acquired data across different style dimensions without any additional annotations.

276, **TITLE:** FUDGE: Controlled Text Generation With Future Discriminators

<https://www.aclweb.org/anthology/2021.naacl-main.276>

**AUTHORS:** Kevin Yang, Dan Klein

**HIGHLIGHT:** We propose Future Discriminators for Generation (FUDGE), a flexible and modular method for controlled text generation.

277, **TITLE:** Controllable Text Simplification with Explicit Paraphrasing

<https://www.aclweb.org/anthology/2021.naacl-main.277>

**AUTHORS:** Mounica Maddela, Fernando Alva-Manchego, Wei Xu

**HIGHLIGHT:** In this paper, we propose a novel hybrid approach that leverages linguistically-motivated rules for splitting and deletion, and couples them with a neural paraphrasing model to produce varied rewriting styles.

278, **TITLE:** Knowledge Graph Based Synthetic Corpus Generation for Knowledge-Enhanced Language Model Pre-training

<https://www.aclweb.org/anthology/2021.naacl-main.278>

**AUTHORS:** Oshin Agarwal, Heming Ge, Siamak Shakeri, Rami Al-Rfou

**HIGHLIGHT:** In this paper, however, we verbalize the entire English Wikidata KG, and discuss the unique challenges associated with a broad, open-domain, large-scale verbalization.

279, **TITLE:** Choose Your Own Adventure: Paired Suggestions in Collaborative Writing for Evaluating Story Generation Models

<https://www.aclweb.org/anthology/2021.naacl-main.279>

**AUTHORS:** Elizabeth Clark, Noah A. Smith

**HIGHLIGHT:** We present Choose Your Own Adventure, a collaborative writing setup for pairwise model evaluation.

280, **TITLE:** InfoXLM: An Information-Theoretic Framework for Cross-Lingual Language Model Pre-Training

<https://www.aclweb.org/anthology/2021.naacl-main.280>

**AUTHORS:** Zewen Chi, Li Dong, Furu Wei, Nan Yang, Saksham Singhal, Wenhui Wang, Xia Song, Xian-Ling Mao, Heyan Huang, Ming Zhou

**HIGHLIGHT:** In this work, we present an information-theoretic framework that formulates cross-lingual language model pre-training as maximizing mutual information between multilingual-multi-granularity texts.

281, **TITLE:** Context-Interactive Pre-Training for Document Machine Translation

<https://www.aclweb.org/anthology/2021.naacl-main.281>

AUTHORS: Pengcheng Yang, Pei Zhang, Boxing Chen, Jun Xie, Weihua Luo  
HIGHLIGHT: To remedy this, here we propose a simple yet effective context-interactive pre-training approach, which targets benefiting from external large-scale corpora.

282, TITLE: Code-Mixing on Sesame Street: Dawn of the Adversarial Polyglots  
<https://www.aclweb.org/anthology/2021.naacl-main.282>  
AUTHORS: Samson Tan, Shafiq Joty  
HIGHLIGHT: Inspired by this phenomenon, we present two strong black-box adversarial attacks (one word-level, one phrase-level) for multilingual models that push their ability to handle code-mixed sentences to the limit.

283, TITLE: X-METRA-ADA: Cross-lingual Meta-Transfer learning Adaptation to Natural Language Understanding and Question Answering  
<https://www.aclweb.org/anthology/2021.naacl-main.283>  
AUTHORS: Meryem M?hamdi, Doo Soon Kim, Franck Dernoncourt, Trung Bui, Xiang Ren, Jonathan May  
HIGHLIGHT: In this work, we propose X-METRA-ADA, a cross-lingual METra-TRAnSfer learning ADAPtation approach for NLU.

284, TITLE: Explicit Alignment Objectives for Multilingual Bidirectional Encoders  
<https://www.aclweb.org/anthology/2021.naacl-main.284>  
AUTHORS: Junjie Hu, Melvin Johnson, Orhan Firat, Aditya Siddhant, Graham Neubig  
HIGHLIGHT: In this paper, we present a new method for learning multilingual encoders, AMBER (Aligned Multilingual Bidirectional EncodeR).

285, TITLE: Cross-lingual Cross-modal Pretraining for Multimodal Retrieval  
<https://www.aclweb.org/anthology/2021.naacl-main.285>  
AUTHORS: Hongliang Fei, Tan Yu, Ping Li  
HIGHLIGHT: This paper proposes a new approach to learn cross-lingual cross-modal representations for matching images and their relevant captions in multiple languages.

286, TITLE: Wikipedia Entities as Rendezvous across Languages: Grounding Multilingual Language Models by Predicting Wikipedia Hyperlinks  
<https://www.aclweb.org/anthology/2021.naacl-main.286>  
AUTHORS: Iacer Calixto, Alessandro Raganato, Tommaso Pasini  
HIGHLIGHT: In this work, we propose a language-independent entity prediction task as an intermediate training procedure to ground word representations on entity semantics and bridge the gap across different languages by means of a shared vocabulary of entities.

287, TITLE: multiPROver: Generating Multiple Proofs for Improved Interpretability in Rule Reasoning  
<https://www.aclweb.org/anthology/2021.naacl-main.287>  
AUTHORS: Swarnadeep Saha, Prateek Yadav, Mohit Bansal  
HIGHLIGHT: We propose two variants of a proof-set generation model, multiPROver.

288, TITLE: Adaptable and Interpretable Neural MemoryOver Symbolic Knowledge  
<https://www.aclweb.org/anthology/2021.naacl-main.288>  
AUTHORS: Pat Verga, Haitian Sun, Livio Baldini Soares, William Cohen  
HIGHLIGHT: To address this problem, we develop a neural LM that includes an interpretable neuro-symbolic KB in the form of a "fact memory".

289, TITLE: CLEVR\_HYP: A Challenge Dataset and Baselines for Visual Question Answering with Hypothetical Actions over Images  
<https://www.aclweb.org/anthology/2021.naacl-main.289>  
AUTHORS: Shailaja Keyur Sampat, Akshay Kumar, Yezhou Yang, Chitta Baral  
HIGHLIGHT: In this paper, we take visual understanding to a higher level where systems are challenged to answer questions that involve mentally simulating the hypothetical consequences of performing specific actions in a given scenario.

290, TITLE: Refining Targeted Syntactic Evaluation of Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.290>  
AUTHORS: Benjamin Newman, Kai-Siang Ang, Julia Gong, John Hewitt  
HIGHLIGHT: We argue that current implementations of TSE do not directly capture either of these goals, and propose new metrics to capture each goal separately.

- 291, TITLE: Universal Adversarial Attacks with Natural Triggers for Text Classification  
<https://www.aclweb.org/anthology/2021.naacl-main.291>  
AUTHORS: Liwei Song, Xinwei Yu, Hsuan-Tung Peng, Karthik Narasimhan  
HIGHLIGHT: We leverage an adversarially regularized autoencoder (ARAE) to generate triggers and propose a gradient-based search that aims to maximize the downstream classifier's prediction loss.
- 292, TITLE: QuadrupletBERT: An Efficient Model For Embedding-Based Large-Scale Retrieval  
<https://www.aclweb.org/anthology/2021.naacl-main.292>  
AUTHORS: Peiyang Liu, Sen Wang, Xi Wang, Wei Ye, Shikun Zhang  
HIGHLIGHT: Considering that pre-trained language models like BERT have achieved great success in a wide variety of NLP tasks, we present a QuadrupletBERT model for effective and efficient retrieval in this paper.
- 293, TITLE: Dynamically Disentangling Social Bias from Task-Oriented Representations with Adversarial Attack  
<https://www.aclweb.org/anthology/2021.naacl-main.293>  
AUTHORS: Liwen Wang, Yuanmeng Yan, Keqing He, Yanan Wu, Weiran Xu  
HIGHLIGHT: In this paper, we propose an adversarial disentangled debiasing model to dynamically decouple social bias attributes from the intermediate representations trained on the main task.
- 294, TITLE: An Empirical Investigation of Bias in the Multimodal Analysis of Financial Earnings Calls  
<https://www.aclweb.org/anthology/2021.naacl-main.294>  
AUTHORS: Ramit Sawhney, Arshiya Aggarwal, Rajiv Ratn Shah  
HIGHLIGHT: In this work, we present the first study to discover the gender bias in multimodal volatility prediction due to gender-sensitive audio features and fewer female executives in earnings calls of one of the world's biggest stock indexes, the S&P 500 index.
- 295, TITLE: Beyond Fair Pay: Ethical Implications of NLP Crowdsourcing  
<https://www.aclweb.org/anthology/2021.naacl-main.295>  
AUTHORS: Boaz Shmueli, Jan Fell, Soumya Ray, Lun-Wei Ku  
HIGHLIGHT: We find that the Final Rule, the common ethical framework used by researchers, did not anticipate the use of online crowdsourcing platforms for data collection, resulting in gaps between the spirit and practice of human-subjects ethics in NLP research.
- 296, TITLE: On Transferability of Bias Mitigation Effects in Language Model Fine-Tuning  
<https://www.aclweb.org/anthology/2021.naacl-main.296>  
AUTHORS: Xisen Jin, Francesco Barbieri, Brendan Kennedy, Aida Mostafazadeh Davani, Leonardo Neves, Xiang Ren  
HIGHLIGHT: We explore the feasibility and benefits of upstream bias mitigation (UBM) for reducing bias on downstream tasks, by first applying bias mitigation to an upstream model through fine-tuning and subsequently using it for downstream fine-tuning.
- 297, TITLE: Case Study: Deontological Ethics in NLP  
<https://www.aclweb.org/anthology/2021.naacl-main.297>  
AUTHORS: Shrimai Prabhumoye, Brendon Boldt, Ruslan Salakhutdinov, Alan W Black  
HIGHLIGHT: In this work, we study one ethical theory, namely deontological ethics, from the perspective of NLP.
- 298, TITLE: Privacy Regularization: Joint Privacy-Utility Optimization in Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.298>  
AUTHORS: Fatemehsadat Mireshghallah, Huseyin Inan, Marcello Hasegawa, Victor R'hle, Taylor Berg-Kirkpatrick, Robert Sim  
HIGHLIGHT: In this work, we introduce two privacy-preserving regularization methods for training language models that enable joint optimization of utility and privacy through (1) the use of a discriminator and (2) the inclusion of a novel triplet-loss term.
- 299, TITLE: On the Impact of Random Seeds on the Fairness of Clinical Classifiers  
<https://www.aclweb.org/anthology/2021.naacl-main.299>  
AUTHORS: Silvio Amir, Jan-Willem van de Meent, Byron Wallace  
HIGHLIGHT: Recent work has shown that fine-tuning large networks is surprisingly sensitive to changes in random seed(s). We explore the implications of this phenomenon for model fairness across demographic groups in clinical prediction tasks over electronic health records (EHR) in MIMIC-III -- the standard dataset in clinical NLP research.
- 300, TITLE: Topic Model or Topic Twaddle? Re-evaluating Semantic Interpretability Measures  
<https://www.aclweb.org/anthology/2021.naacl-main.300>

AUTHORS: Caitlin Doogan, Wray Buntine  
HIGHLIGHT: In this paper, we probe the issue of validity in topic model evaluation and assess how informative coherence measures are for specialized collections used in an applied setting.

301, TITLE: Discourse Probing of Pretrained Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.301>  
AUTHORS: Fajri Koto, Jey Han Lau, Timothy Baldwin  
HIGHLIGHT: In this paper, we introduce document-level discourse probing to evaluate the ability of pretrained LMs to capture document-level relations.

302, TITLE: UniDrop: A Simple yet Effective Technique to Improve Transformer without Extra Cost  
<https://www.aclweb.org/anthology/2021.naacl-main.302>  
AUTHORS: Zhen Wu, Lijun Wu, Qi Meng, Yingce Xia, Shufang Xie, Tao Qin, Xinyu Dai, Tie-Yan Liu  
HIGHLIGHT: Therefore, in this paper, we integrate different dropout techniques into the training of Transformer models.

303, TITLE: tWT-WT: A Dataset to Assert the Role of Target Entities for Detecting Stance of Tweets  
<https://www.aclweb.org/anthology/2021.naacl-main.303>  
AUTHORS: Ayush Kaushal, Avirup Saha, Niloy Ganguly  
HIGHLIGHT: Though the task involves reasoning of the tweet with respect to a target, we find that it is possible to achieve high accuracy on several publicly available Twitter stance detection datasets without looking at the target sentence.

304, TITLE: Learning to Learn to be Right for the Right Reasons  
<https://www.aclweb.org/anthology/2021.naacl-main.304>  
AUTHORS: Pride Kavumba, Benjamin Heinzerling, Ana Brassard, Kentaro Inui  
HIGHLIGHT: Here, we propose to explicitly learn a model that does well on both the easy test set with superficial cues and the hard test set without superficial cues.

305, TITLE: Double Perturbation: On the Robustness of Robustness and Counterfactual Bias Evaluation  
<https://www.aclweb.org/anthology/2021.naacl-main.305>  
AUTHORS: Chong Zhang, Jieyu Zhao, Huan Zhang, Kai-Wei Chang, Cho-Jui Hsieh  
HIGHLIGHT: In this paper, we propose a "double perturbation" framework to uncover model weaknesses beyond the test dataset.

306, TITLE: Explaining Neural Network Predictions on Sentence Pairs via Learning Word-Group Masks  
<https://www.aclweb.org/anthology/2021.naacl-main.306>  
AUTHORS: Hanjie Chen, Song Feng, Jatin Ganhotra, Hui Wan, Chulaka Gunasekara, Sachindra Joshi, Yangfeng Ji  
HIGHLIGHT: In this work, we propose the Group Mask (GMASK) method to implicitly detect word correlations by grouping correlated words from the input text pair together and measure their contribution to the corresponding NLP tasks as a whole.

307, TITLE: Almost Free Semantic Draft for Neural Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.307>  
AUTHORS: Xi Ai, Bin Fang  
HIGHLIGHT: In this work, to inject global information but also save cost, we present an efficient method to sample and consider a semantic draft as global information from semantic space for decoding with almost free of cost.

308, TITLE: Pruning-then-Expanding Model for Domain Adaptation of Neural Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.308>  
AUTHORS: Shuhao Gu, Yang Feng, Wanying Xie  
HIGHLIGHT: To address these three problems, we propose a method of "divide and conquer" which is based on the importance of neurons or parameters for the translation model.

309, TITLE: Multi-Hop Transformer for Document-Level Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.309>  
AUTHORS: Long Zhang, Tong Zhang, Haibo Zhang, Baosong Yang, Wei Ye, Shikun Zhang  
HIGHLIGHT: To this end, we propose a novel Multi-Hop Transformer (MHT) which offers NMT abilities to explicitly model the human-like draft-editing and reasoning process.

310, TITLE: Continual Learning for Neural Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.310>  
AUTHORS: Yue Cao, Hao-Ran Wei, Boxing Chen, Xiaojun Wan



- HIGHLIGHT:** In this work, we propose a new continual learning framework for NMT models.
- 311, **TITLE:** Self-Training for Unsupervised Neural Machine Translation in Unbalanced Training Data Scenarios  
<https://www.aclweb.org/anthology/2021.naacl-main.311>  
**AUTHORS:** Haipeng Sun, Rui Wang, Kehai Chen, Masao Utiyama, Eiichiro Sumita, Tiejun Zhao  
**HIGHLIGHT:** Based on this scenario, we propose UNMT self-training mechanisms to train a robust UNMT system and improve its performance in this case.
- 312, **TITLE:** Smart-Start Decoding for Neural Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.312>  
**AUTHORS:** Jian Yang, Shuming Ma, Dongdong Zhang, Juncheng Wan, Zhoujun Li, Ming Zhou  
**HIGHLIGHT:** In this work, we propose a novel method that breaks up the limitation of these decoding orders, called Smart-Start decoding.
- 313, **TITLE:** Multi-Task Learning with Shared Encoder for Non-Autoregressive Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.313>  
**AUTHORS:** Yongchang Hao, Shilin He, Wenxiang Jiao, Zhaopeng Tu, Michael Lyu, Xing Wang  
**HIGHLIGHT:** In this work, we hypothesize and empirically verify that AT and NAT encoders capture different linguistic properties of source sentences.
- 314, **TITLE:** ER-AE: Differentially Private Text Generation for Authorship Anonymization  
<https://www.aclweb.org/anthology/2021.naacl-main.314>  
**AUTHORS:** Haohan Bo, Steven H. H. Ding, Benjamin C. M. Fung, Farkhund Iqbal  
**HIGHLIGHT:** We propose a novel text generation model with a two-set exponential mechanism for authorship anonymization.
- 315, **TITLE:** Distantly Supervised Transformers For E-Commerce Product QA  
<https://www.aclweb.org/anthology/2021.naacl-main.315>  
**AUTHORS:** Happy Mittal, Aniket Chakrabarti, Belhassen Bayar, Animesh Anant Sharma, Nikhil Rasiwasia  
**HIGHLIGHT:** We propose a practical instant question answering (QA) system on product pages of e-commerce services, where for each user query, relevant community question answer (CQA) pairs are retrieved.
- 316, **TITLE:** Quantitative Day Trading from Natural Language using Reinforcement Learning  
<https://www.aclweb.org/anthology/2021.naacl-main.316>  
**AUTHORS:** Ramit Sawhney, Arnav Wadhwa, Shivam Agarwal, Rajiv Ratn Shah  
**HIGHLIGHT:** Building on these shortcomings, we propose a deep reinforcement learning approach that makes time-aware decisions to trade stocks while optimizing profit using textual data.
- 317, **TITLE:** Restoring and Mining the Records of the Joseon Dynasty via Neural Language Modeling and Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.317>  
**AUTHORS:** Kyeongpil Kang, Kyohoon Jin, Soyoung Yang, Soojin Jang, Jaegul Choo, Youngbin Kim  
**HIGHLIGHT:** In response, we present a multi-task learning approach to restore and translate historical documents based on a self-attention mechanism, specifically utilizing two Korean historical records, ones of the most voluminous historical records in the world.
- 318, **TITLE:** Modeling Diagnostic Label Correlation for Automatic ICD Coding  
<https://www.aclweb.org/anthology/2021.naacl-main.318>  
**AUTHORS:** Shang-Chi Tsai, Chao-Wei Huang, Yun-Nung Chen  
**HIGHLIGHT:** To address this problem, we propose a two-stage framework to improve automatic ICD coding by capturing the label correlation.
- 319, **TITLE:** Self-Supervised Contrastive Learning for Efficient User Satisfaction Prediction in Conversational Agents  
<https://www.aclweb.org/anthology/2021.naacl-main.319>  
**AUTHORS:** Mohammad Kachuee, Hao Yuan, Young-Bum Kim, Sungjin Lee  
**HIGHLIGHT:** In this paper, we suggest a self-supervised contrastive learning approach that leverages the pool of unlabeled data to learn user-agent interactions.
- 320, **TITLE:** A recipe for annotating grounded clarifications  
<https://www.aclweb.org/anthology/2021.naacl-main.320>  
**AUTHORS:** Luciana Benotti, Patrick Blackburn

**HIGHLIGHT:** In this paper, we argue that dialogue clarification mechanisms make explicit the process of interpreting the communicative intents of the speaker's utterances by grounding them in the various modalities in which the dialogue is situated.

321, **TITLE:** Grey-box Adversarial Attack And Defence For Sentiment Classification  
<https://www.aclweb.org/anthology/2021.naacl-main.321>  
**AUTHORS:** Ying Xu, Xu Zhong, Antonio Jimeno Yepes, Jey Han Lau  
**HIGHLIGHT:** We introduce a grey-box adversarial attack and defence framework for sentiment classification.

322, **TITLE:** How low is too low? A monolingual take on lemmatisation in Indian languages  
<https://www.aclweb.org/anthology/2021.naacl-main.322>  
**AUTHORS:** Kumar Saunack, Kumar Saurav, Pushpak Bhattacharyya  
**HIGHLIGHT:** In this paper, we devote our attention to lemmatisation for low resource, morphologically rich scheduled Indian languages using neural methods.

323, **TITLE:** Causal Effects of Linguistic Properties  
<https://www.aclweb.org/anthology/2021.naacl-main.323>  
**AUTHORS:** Reid Pryzant, Dallas Card, Dan Jurafsky, Victor Veitch, Dhanya Sridhar  
**HIGHLIGHT:** Based on these results, we introduce TextCause, an algorithm for estimating causal effects of linguistic properties.

324, **TITLE:** Dynabench: Rethinking Benchmarking in NLP  
<https://www.aclweb.org/anthology/2021.naacl-main.324>  
**AUTHORS:** Douwe Kiela, Max Bartolo, Yixin Nie, Divyansh Kaushik, Atticus Geiger, Zhengxuan Wu, Bertie Vidgen, Grusha Prasad, Amanpreet Singh, Pratik Ringshia, Zhiyi Ma, Tristan Thrush, Sebastian Riedel, Zeerak Waseem, Pontus Stenetorp, Robin Jia, Mohit Bansal, Christopher Potts, Adina Williams  
**HIGHLIGHT:** In this paper, we argue that Dynabench addresses a critical need in our community: contemporary models quickly achieve outstanding performance on benchmark tasks but nonetheless fail on simple challenge examples and falter in real-world scenarios.

325, **TITLE:** Translational NLP: A New Paradigm and General Principles for Natural Language Processing Research  
<https://www.aclweb.org/anthology/2021.naacl-main.325>  
**AUTHORS:** Denis Newman-Griffis, Jill Fain Lehman, Carolyn Ros?, Harry Hochheiser  
**HIGHLIGHT:** We show that many significant advances in NLP research have emerged from the intersection of basic principles with application needs, and present a conceptual framework outlining the stakeholders and key questions in translational research.

326, **TITLE:** Predicting Discourse Trees from Transformer-based Neural Summarizers  
<https://www.aclweb.org/anthology/2021.naacl-main.326>  
**AUTHORS:** Wen Xiao, Patrick Huber, Giuseppe Carenini  
**HIGHLIGHT:** In this paper, we explore whether this synergy between discourse and summarization is bidirectional, by inferring document-level discourse trees from pre-trained neural summarizers.

327, **TITLE:** Probing for Bridging Inference in Transformer Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.327>  
**AUTHORS:** Onkar Pandit, Yufang Hou  
**HIGHLIGHT:** We first investigate individual attention heads in BERT and observe that attention heads at higher layers prominently focus on bridging relations in-comparison with the lower and middle layers, also, few specific attention heads concentrate consistently on bridging.

328, **TITLE:** Is Incoherence Surprising? Targeted Evaluation of Coherence Prediction from Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.328>  
**AUTHORS:** Anne Beyer, Sharid Lo?iciga, David Schlangen  
**HIGHLIGHT:** Extending the targeted evaluation paradigm for neural language models (Marvin and Linzen, 2018) to phenomena beyond syntax, we show that this paradigm is equally suited to evaluate linguistic qualities that contribute to the notion of coherence.

329, **TITLE:** Stay Together: A System for Single and Split-antecedent Anaphora Resolution  
<https://www.aclweb.org/anthology/2021.naacl-main.329>  
**AUTHORS:** Juntao Yu, Nafise Sadat Moosavi, Silviu Paun, Massimo Poesio  
**HIGHLIGHT:** In this work, we introduce a system that resolves both single and split-antecedent anaphors, and evaluate it in a more realistic setting that uses predicted mentions.

- 330, TITLE: Redefining Absent Keyphrases and their Effect on Retrieval Effectiveness  
<https://www.aclweb.org/anthology/2021.naacl-main.330>  
AUTHORS: Florian Boudin, Ygor Gallina  
HIGHLIGHT: In this paper, we discuss the usefulness of absent keyphrases from an Information Retrieval (IR) perspective, and show that the commonly drawn distinction between present and absent keyphrases is not made explicit enough.
- 331, TITLE: CoRT: Complementary Rankings from Transformers  
<https://www.aclweb.org/anthology/2021.naacl-main.331>  
AUTHORS: Marco Wrzalik, Dirk Krechel  
HIGHLIGHT: In this context we propose CoRT, a simple neural first-stage ranking model that leverages contextual representations from pretrained language models such as BERT to complement term-based ranking functions while causing no significant delay at query time.
- 332, TITLE: Multi-source Neural Topic Modeling in Multi-view Embedding Spaces  
<https://www.aclweb.org/anthology/2021.naacl-main.332>  
AUTHORS: Pankaj Gupta, Yatin Chaudhary, Hinrich Sch?tze  
HIGHLIGHT: This work presents a novel neural topic modeling framework using multi-view embedding spaces: (1) pretrained topic-embeddings, and (2) pretrained word-embeddings (context-insensitive from Glove and context-sensitive from BERT models) jointly from one or many sources to improve topic quality and better deal with polysemy.
- 333, TITLE: Inductive Topic Variational Graph Auto-Encoder for Text Classification  
<https://www.aclweb.org/anthology/2021.naacl-main.333>  
AUTHORS: Qianqian Xie, Jimin Huang, Pan Du, Min Peng, Jian-Yun Nie  
HIGHLIGHT: To address these issues, we propose a novel model named inductive Topic Variational Graph Auto-Encoder (T-VGAE), which incorporates a topic model into variational graph-auto-encoder (VGAE) to capture the hidden semantic information between documents and words.
- 334, TITLE: Self-Alignment Pretraining for Biomedical Entity Representations  
<https://www.aclweb.org/anthology/2021.naacl-main.334>  
AUTHORS: Fangyu Liu, Ehsan Shareghi, Zaiqiao Meng, Marco Basaldella, Nigel Collier  
HIGHLIGHT: To address this challenge, we propose SapBERT, a pretraining scheme that self-aligns the representation space of biomedical entities.
- 335, TITLE: TaxoClass: Hierarchical Multi-Label Text Classification Using Only Class Names  
<https://www.aclweb.org/anthology/2021.naacl-main.335>  
AUTHORS: Jiaming Shen, Wenda Qiu, Yu Meng, Jingbo Shang, Xiang Ren, Jiawei Han  
HIGHLIGHT: In this paper, we explore to conduct HMTC based on only class surface names as supervision signals.
- 336, TITLE: MERMAID: Metaphor Generation with Symbolism and Discriminative Decoding  
<https://www.aclweb.org/anthology/2021.naacl-main.336>  
AUTHORS: Tuhin Chakrabarty, Xurui Zhang, Smaranda Muresan, Nanyun Peng  
HIGHLIGHT: In this paper, we aim to generate a metaphoric sentence given a literal expression by replacing relevant verbs.
- 337, TITLE: On Learning Text Style Transfer with Direct Rewards  
<https://www.aclweb.org/anthology/2021.naacl-main.337>  
AUTHORS: Yixin Liu, Graham Neubig, John Wieting  
HIGHLIGHT: In this paper, we explore training algorithms that instead optimize reward functions that explicitly consider different aspects of the style-transferred outputs.
- 338, TITLE: Focused Attention Improves Document-Grounded Generation  
<https://www.aclweb.org/anthology/2021.naacl-main.338>  
AUTHORS: Shrimai Prabhumoye, Kazuma Hashimoto, Yingbo Zhou, Alan W Black, Ruslan Salakhutdinov  
HIGHLIGHT: Our work introduces two novel adaptations of large scale pre-trained encoder-decoder models focusing on building context driven representation of the document and enabling specific attention to the information in the document.
- 339, TITLE: NeuroLogic Decoding: (Un)supervised Neural Text Generation with Predicate Logic Constraints  
<https://www.aclweb.org/anthology/2021.naacl-main.339>  
AUTHORS: Ximing Lu, Peter West, Rowan Zellers, Ronan Le Bras, Chandra Bhagavatula, Yejin Choi

**HIGHLIGHT:** We propose NeuroLogic Decoding, a simple yet effective algorithm that enables neural language models - supervised or not - to generate fluent text while satisfying complex lexical constraints.

340, **TITLE:** Ask what 's missing and what 's useful: Improving Clarification Question Generation using Global Knowledge  
<https://www.aclweb.org/anthology/2021.naacl-main.340>

**AUTHORS:** Bodhisattwa Prasad Majumder, Sudha Rao, Michel Galley, Julian McAuley

**HIGHLIGHT:** Inspired by this, we propose a model for clarification question generation where we first identify what is missing by taking a difference between the global and the local view and then train a model to identify what is useful and generate a question about it.

341, **TITLE:** Progressive Generation of Long Text with Pretrained Language Models

<https://www.aclweb.org/anthology/2021.naacl-main.341>

**AUTHORS:** Bowen Tan, Zichao Yang, Maruan Al-Shedivat, Eric Xing, Zhiting Hu

**HIGHLIGHT:** To overcome the limitations, we propose a simple but effective method of generating text in a progressive manner, inspired by generating images from low to high resolution.

342, **TITLE:** SOCCER: An Information-Sparse Discourse State Tracking Collection in the Sports Commentary Domain

<https://www.aclweb.org/anthology/2021.naacl-main.342>

**AUTHORS:** Ruo Chen Zhang, Carsten Eickhoff

**HIGHLIGHT:** This paper proposes to turn to simplified, fully observable systems that show some of these properties: Sports events.

343, **TITLE:** Plot-guided Adversarial Example Construction for Evaluating Open-domain Story Generation

<https://www.aclweb.org/anthology/2021.naacl-main.343>

**AUTHORS:** Sarik Ghazarian, Zixi Liu, Akash S M, Ralph Weischedel, Aram Galstyan, Nanyun Peng

**HIGHLIGHT:** We propose to tackle these issues by generating a more comprehensive set of implausible stories using plots, which are structured representations of controllable factors used to generate stories.

344, **TITLE:** MultiOpEd: A Corpus of Multi-Perspective News Editorials

<https://www.aclweb.org/anthology/2021.naacl-main.344>

**AUTHORS:** Siyi Liu, Sihao Chen, Xander Uyttendaele, Dan Roth

**HIGHLIGHT:** We propose MultiOpEd, an open-domain news editorial corpus that supports various tasks pertaining to the argumentation structure in news editorials, focusing on automatic perspective discovery.

345, **TITLE:** Swords: A Benchmark for Lexical Substitution with Improved Data Coverage and Quality

<https://www.aclweb.org/anthology/2021.naacl-main.345>

**AUTHORS:** Mina Lee, Chris Donahue, Robin Jia, Alexander Iyabor, Percy Liang

**HIGHLIGHT:** We release a new benchmark for lexical substitution, the task of finding appropriate substitutes for a target word in a context.

346, **TITLE:** "I'm Not Mad": Commonsense Implications of Negation and Contradiction

<https://www.aclweb.org/anthology/2021.naacl-main.346>

**AUTHORS:** Liwei Jiang, Antoine Bosselut, Chandra Bhagavatula, Yejin Choi

**HIGHLIGHT:** In this paper, we present the first comprehensive study focusing on commonsense implications of negated statements and contradictions.

347, **TITLE:** Identifying Medical Self-Disclosure in Online Communities

<https://www.aclweb.org/anthology/2021.naacl-main.347>

**AUTHORS:** Mina Valizadeh, Pardis Ranjbar-Noiey, Cornelia Caragea, Natalie Parde

**HIGHLIGHT:** We address this shortcoming by introducing a new dataset of health-related posts collected from online social platforms, categorized into three groups (No Self-Disclosure, Possible Self-Disclosure, and Clear Self-Disclosure) with high inter-annotator agreement ( $\kappa=0.88$ ).

348, **TITLE:** Language in a (Search) Box: Grounding Language Learning in Real-World Human-Machine Interaction

<https://www.aclweb.org/anthology/2021.naacl-main.348>

**AUTHORS:** Federico Bianchi, Ciro Greco, Jacopo Tagliabue

**HIGHLIGHT:** We investigate grounded language learning through real-world data, by modelling a teacher-learner dynamics through the natural interactions occurring between users and search engines; in particular, we explore the emergence of semantic generalization from unsupervised dense representations outside of synthetic environments.

- 349, TITLE: Finding Concept-specific Biases in Form?Meaning Associations  
<https://www.aclweb.org/anthology/2021.naacl-main.349>  
AUTHORS: Tiago Pimentel, Brian Roark, S?ren Wichmann, Ryan Cotterell, Dami?n Blasi  
HIGHLIGHT: This work presents an information-theoretic operationalisation of cross-linguistic non-arbitrariness.
- 350, TITLE: How (Non-)Optimal is the Lexicon?  
<https://www.aclweb.org/anthology/2021.naacl-main.350>  
AUTHORS: Tiago Pimentel, Irene Nikkarinen, Kyle Mahowald, Ryan Cotterell, Dami?n Blasi  
HIGHLIGHT: Taking a coding-theoretic view of the lexicon and making use of a novel generative statistical model, we define upper bounds for the compressibility of the lexicon under various constraints.
- 351, TITLE: Word Complexity is in the Eye of the Beholder  
<https://www.aclweb.org/anthology/2021.naacl-main.351>  
AUTHORS: Sian Gooding, Ekaterina Kochmar, Seid Muhie Yimam, Chris Biemann  
HIGHLIGHT: In this paper, we investigate which aspects contribute to the notion of lexical complexity in various groups of readers, focusing on native and non-native speakers of English, and how the notion of complexity changes depending on the proficiency level of a non-native reader.
- 352, TITLE: Linguistic Complexity Loss in Text-Based Therapy  
<https://www.aclweb.org/anthology/2021.naacl-main.352>  
AUTHORS: Jason Wei, Kelly Finn, Emma Templeton, Thalia Wheatley, Soroush Vosoughi  
HIGHLIGHT: In this paper, we analyze linguistic complexity correlates of mental health in the online therapy messages sent between therapists and 7,170 clients who provided 30,437 corresponding survey responses on their anxiety.
- 353, TITLE: Ab Antiquo: Neural Proto-language Reconstruction  
<https://www.aclweb.org/anthology/2021.naacl-main.353>  
AUTHORS: Carlo Meloni, Shauli Ravfogel, Yoav Goldberg  
HIGHLIGHT: We address the task of proto-word reconstruction, in which the model is exposed to cognates in contemporary daughter languages, and has to predict the proto word in the ancestor language.
- 354, TITLE: On Biasing Transformer Attention Towards Monotonicity  
<https://www.aclweb.org/anthology/2021.naacl-main.354>  
AUTHORS: Annette Rios, Chantal Amrhein, No?mi Aepli, Rico Sennrich  
HIGHLIGHT: In this work, we introduce a monotonicity loss function that is compatible with standard attention mechanisms and test it on several sequence-to-sequence tasks: grapheme-to-phoneme conversion, morphological inflection, transliteration, and dialect normalization.
- 355, TITLE: Extracting a Knowledge Base of Mechanisms from COVID-19 Papers  
<https://www.aclweb.org/anthology/2021.naacl-main.355>  
AUTHORS: Tom Hope, Aida Amini, David Wadden, Madeleine van Zuylen, Sravanthi Parasa, Eric Horvitz, Daniel Weld, Roy Schwartz, Hannaneh Hajishirzi  
HIGHLIGHT: We pursue the construction of a knowledge base (KB) of mechanisms—a fundamental concept across the sciences, which encompasses activities, functions and causal relations, ranging from cellular processes to economic impacts.
- 356, TITLE: Constrained Multi-Task Learning for Event Coreference Resolution  
<https://www.aclweb.org/anthology/2021.naacl-main.356>  
AUTHORS: Jing Lu, Vincent Ng  
HIGHLIGHT: We propose a neural event coreference model in which event coreference is jointly trained with five tasks: trigger detection, entity coreference, anaphoricity determination, realis detection, and argument extraction.
- 357, TITLE: Empirical Evaluation of Pre-trained Transformers for Human-Level NLP: The Role of Sample Size and Dimensionality  
<https://www.aclweb.org/anthology/2021.naacl-main.357>  
AUTHORS: Adithya V Ganesan, Matthew Matero, Aravind Reddy Ravula, Huy Vu, H. Andrew Schwartz  
HIGHLIGHT: Here, we provide a systematic study on the role of dimension reduction methods (principal components analysis, factorization techniques, or multi-layer auto-encoders) as well as the dimensionality of embedding vectors and sample sizes as a function of predictive performance.
- 358, TITLE: Leveraging Deep Representations of Radiology Reports in Survival Analysis for Predicting Heart Failure Patient Mortality  
<https://www.aclweb.org/anthology/2021.naacl-main.358>

AUTHORS: Hyun Gi Lee, Evan Sholle, Ashley Beecy, Subhi Al-Aref, Yifan Peng  
HIGHLIGHT: In this work, we present a novel method of using BERT-based hidden layer representations of clinical texts as covariates for proportional hazards models to predict patient survival outcomes.

359, TITLE: On the Use of Context for Predicting Citation Worthiness of Sentences in Scholarly Articles  
<https://www.aclweb.org/anthology/2021.naacl-main.359>  
AUTHORS: Rakesh Gosangi, Ravneet Arora, Mohsen Gheisarieha, Debanjan Mahata, Haimin Zhang  
HIGHLIGHT: In this paper, we study the importance of context in predicting the citation worthiness of sentences in scholarly articles.

360, TITLE: Data and Model Distillation as a Solution for Domain-transferable Fact Verification  
<https://www.aclweb.org/anthology/2021.naacl-main.360>  
AUTHORS: Mitch Paul Mithun, Sandeep Surtwal, Mihai Surdeanu  
HIGHLIGHT: While neural networks produce state-of-the-art performance in several NLP tasks, they generally depend heavily on lexicalized information, which transfer poorly between domains. We present a combination of two strategies to mitigate this dependence on lexicalized information in fact verification tasks

361, TITLE: Adapting Coreference Resolution for Processing Violent Death Narratives  
<https://www.aclweb.org/anthology/2021.naacl-main.361>  
AUTHORS: Ankit Uppunda, Susan Cochran, Jacob Foster, Alina Arseniev-Koehler, Vickie Mays, Kai-Wei Chang  
HIGHLIGHT: In this paper, we analyzed the challenges of coreference resolution in an exemplary form of administrative text written in English: violent death narratives from the USA's Centers for Disease Control's (CDC) National Violent Death Reporting System.

362, TITLE: Time-Stamped Language Model: Teaching Language Models to Understand The Flow of Events  
<https://www.aclweb.org/anthology/2021.naacl-main.362>  
AUTHORS: Hossein Rajaby Faghihi, Parisa Kordjamshidi  
HIGHLIGHT: Tracking entities throughout a procedure described in a text is challenging due to the dynamic nature of the world described in the process. Firstly, we propose to formulate this task as a question answering problem.

363, TITLE: If You Want to Go Far Go Together: Unsupervised Joint Candidate Evidence Retrieval for Multi-hop Question Answering  
<https://www.aclweb.org/anthology/2021.naacl-main.363>  
AUTHORS: Vikas Yadav, Steven Bethard, Mihai Surdeanu  
HIGHLIGHT: To retrieve such facts, we propose a simple approach that retrieves and reranks set of evidence facts jointly.

364, TITLE: SPARTQA: A Textual Question Answering Benchmark for Spatial Reasoning  
<https://www.aclweb.org/anthology/2021.naacl-main.364>  
AUTHORS: Roshanak Mirzaee, Hossein Rajaby Faghihi, Qiang Ning, Parisa Kordjamshidi  
HIGHLIGHT: This paper proposes a question-answering (QA) benchmark for spatial reasoning on natural language text which contains more realistic spatial phenomena not covered by prior work and is challenging for state-of-the-art language models (LM).

365, TITLE: A Dataset of Information-Seeking Questions and Answers Anchored in Research Papers  
<https://www.aclweb.org/anthology/2021.naacl-main.365>  
AUTHORS: Pradeep Dasigi, Kyle Lo, Iz Beltagy, Arman Cohan, Noah A. Smith, Matt Gardner  
HIGHLIGHT: We therefore present Qasper, a dataset of 5049 questions over 1585 Natural Language Processing papers.

366, TITLE: Differentiable Open-Ended Commonsense Reasoning  
<https://www.aclweb.org/anthology/2021.naacl-main.366>  
AUTHORS: Bill Yuchen Lin, Haitian Sun, Bhuwan Dhingra, Manzil Zaheer, Xiang Ren, William Cohen  
HIGHLIGHT: As an approach to OpenCSR, we propose DrFact, an efficient Differentiable model for multi-hop Reasoning over knowledge Facts.

367, TITLE: Does Structure Matter? Encoding Documents for Machine Reading Comprehension  
<https://www.aclweb.org/anthology/2021.naacl-main.367>  
AUTHORS: Hui Wan, Song Feng, Chulaka Gunasekara, Siva Sankalp Patel, Sachindra Joshi, Luis Lastras  
HIGHLIGHT: This work proposes a new Transformer-based method that reads a document as tree slices.

368, TITLE: Multi-Step Reasoning Over Unstructured Text with Beam Dense Retrieval  
<https://www.aclweb.org/anthology/2021.naacl-main.368>

AUTHORS: Chen Zhao, Chenyan Xiong, Jordan Boyd-Graber, Hal Daum? III  
HIGHLIGHT: Building on dense retrieval methods, we propose a new multi-step retrieval approach (BeamDR) that iteratively forms an evidence chain through beam search in dense representations.

369, TITLE: Scalable and Interpretable Semantic Change Detection  
<https://www.aclweb.org/anthology/2021.naacl-main.369>  
AUTHORS: Syrielle Montariol, Matej Martinc, Lidia Pivovarov  
HIGHLIGHT: We propose a novel scalable method for word usage-change detection that offers large gains in processing time and significant memory savings while offering the same interpretability and better performance than unscalable methods.

370, TITLE: Scalar Adjective Identification and Multilingual Ranking  
<https://www.aclweb.org/anthology/2021.naacl-main.370>  
AUTHORS: Aina Gar? Soler, Marianna Apidianaki  
HIGHLIGHT: We introduce a new multilingual dataset in order to promote research on scalar adjectives in new languages.

371, TITLE: ESC: Redesigning WSD with Extractive Sense Comprehension  
<https://www.aclweb.org/anthology/2021.naacl-main.371>  
AUTHORS: Edoardo Barba, Tommaso Pasini, Roberto Navigli  
HIGHLIGHT: We cope with this issue by reframing WSD as a span extraction problem - which we called Extractive Sense Comprehension (ESC) - and propose ESCHER, a transformer-based neural architecture for this new formulation.

372, TITLE: Recent advances in neural metaphor processing: A linguistic, cognitive and social perspective  
<https://www.aclweb.org/anthology/2021.naacl-main.372>  
AUTHORS: Xiaoyu Tong, Ekaterina Shutova, Martha Lewis  
HIGHLIGHT: This paper provides a comprehensive review and discussion of recent developments in automated metaphor processing, in light of the findings about metaphor in the mind, language, and communication, and from the perspective of downstream NLP tasks.

373, TITLE: Constructing Taxonomies from Pretrained Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.373>  
AUTHORS: Catherine Chen, Kevin Lin, Dan Klein  
HIGHLIGHT: We present a method for constructing taxonomic trees (e.g., WordNet) using pretrained language models.

374, TITLE: Event Representation with Sequential, Semi-Supervised Discrete Variables  
<https://www.aclweb.org/anthology/2021.naacl-main.374>  
AUTHORS: Mehdi Rezaee, Francis Ferraro  
HIGHLIGHT: Within the context of event modeling and understanding, we propose a new method for neural sequence modeling that takes partially-observed sequences of discrete, external knowledge into account.

375, TITLE: Seq2Emo: A Sequence to Multi-Label Emotion Classification Model  
<https://www.aclweb.org/anthology/2021.naacl-main.375>  
AUTHORS: Chenyang Huang, Amine Trabelsi, Xuebin Qin, Nawshad Farruque, Lili Mou, Osmar Za?ane  
HIGHLIGHT: In this work, we propose a sequence-to-emotion (Seq2Emo) approach, which implicitly models emotion correlations in a bi-directional decoder.

376, TITLE: Knowledge Enhanced Masked Language Model for Stance Detection  
<https://www.aclweb.org/anthology/2021.naacl-main.376>  
AUTHORS: Kornraphop Kawintiranon, Lisa Singh  
HIGHLIGHT: In this paper, we propose a novel BERT-based fine-tuning method that enhances the masked language model for stance detection.

377, TITLE: Learning Paralinguistic Features from Audiobooks through Style Voice Conversion  
<https://www.aclweb.org/anthology/2021.naacl-main.377>  
AUTHORS: Zakaria Aldeneh, Matthew Perez, Emily Mower Provost  
HIGHLIGHT: In this work, we present a new framework that enables a neural network to learn to extract paralinguistic attributes from speech using data that are not annotated for emotion.

378, TITLE: Adapting BERT for Continual Learning of a Sequence of Aspect Sentiment Classification Tasks  
<https://www.aclweb.org/anthology/2021.naacl-main.378>  
AUTHORS: Zixuan Ke, Hu Xu, Bing Liu

**HIGHLIGHT:** This paper proposes a novel capsule network based model called B-CL to address these issues.

379, **TITLE:** Adversarial Learning for Zero-Shot Stance Detection on Social Media

<https://www.aclweb.org/anthology/2021.naacl-main.379>

**AUTHORS:** Emily Allaway, Malavika Srikanth, Kathleen McKeown

**HIGHLIGHT:** In this work, we propose a new model for zero-shot stance detection on Twitter that uses adversarial learning to generalize across topics.

380, **TITLE:** Efficiently Summarizing Text and Graph Encodings of Multi-Document Clusters

<https://www.aclweb.org/anthology/2021.naacl-main.380>

**AUTHORS:** Ramakanth Pasunuru, Mengwen Liu, Mohit Bansal, Sujith Ravi, Markus Dreyer

**HIGHLIGHT:** This paper presents an efficient graph-enhanced approach to multi-document summarization (MDS) with an encoder-decoder Transformer model.

381, **TITLE:** Enriching Transformers with Structured Tensor-Product Representations for Abstractive Summarization

<https://www.aclweb.org/anthology/2021.naacl-main.381>

**AUTHORS:** Yichen Jiang, Asli Celikyilmaz, Paul Smolensky, Paul Soulos, Sudha Rao, Hamid Palangi, Roland Fernandez, Caitlin Smith, Mohit Bansal, Jianfeng Gao

**HIGHLIGHT:** In this paper, we adapt TP-Transformer (Schlag et al., 2019), an architecture that enriches the original Transformer (Vaswani et al., 2017) with the explicitly compositional Tensor Product Representation (TPR), for the task of abstractive summarization.

382, **TITLE:** What's in a Summary? Laying the Groundwork for Advances in Hospital-Course Summarization

<https://www.aclweb.org/anthology/2021.naacl-main.382>

**AUTHORS:** Griffin Adams, Emily Alsentzer, Mert Ketenci, Jason Zucker, No?mie Elhadad

**HIGHLIGHT:** Here, we introduce the task of hospital-course summarization.

383, **TITLE:** Understanding Factuality in Abstractive Summarization with FRANK: A Benchmark for Factuality Metrics

<https://www.aclweb.org/anthology/2021.naacl-main.383>

**AUTHORS:** Artidoro Pagnoni, Vidhisha Balachandran, Yulia Tsvetkov

**HIGHLIGHT:** To address these limitations, we devise a typology of factual errors and use it to collect human annotations of generated summaries from state-of-the-art summarization systems for the CNN/DM and XSum datasets.

384, **TITLE:** GSum: A General Framework for Guided Neural Abstractive Summarization

<https://www.aclweb.org/anthology/2021.naacl-main.384>

**AUTHORS:** Zi-Yi Dou, Pengfei Liu, Hiroaki Hayashi, Zhengbao Jiang, Graham Neubig

**HIGHLIGHT:** In this paper, we propose a general and extensible guided summarization framework (GSum) that can effectively take different kinds of external guidance as input, and we perform experiments across several different varieties.

385, **TITLE:** What Will it Take to Fix Benchmarking in Natural Language Understanding?

<https://www.aclweb.org/anthology/2021.naacl-main.385>

**AUTHORS:** Samuel R. Bowman, George Dahl

**HIGHLIGHT:** In this position paper, we lay out four criteria that we argue NLU benchmarks should meet.

386, **TITLE:** TuringAdvice: A Generative and Dynamic Evaluation of Language Use

<https://www.aclweb.org/anthology/2021.naacl-main.386>

**AUTHORS:** Rowan Zellers, Ari Holtzman, Elizabeth Clark, Lianhui Qin, Ali Farhadi, Yejin Choi

**HIGHLIGHT:** We propose TuringAdvice, a new challenge task and dataset for language understanding models.

387, **TITLE:** Multitask Learning for Emotionally Analyzing Sexual Abuse Disclosures

<https://www.aclweb.org/anthology/2021.naacl-main.387>

**AUTHORS:** Ramit Sawhney, Puneet Mathur, Taru Jain, Akash Kumar Gautam, Rajiv Ratn Shah

**HIGHLIGHT:** We formulate the task of identifying narratives related to the sexual abuse disclosures in online posts as a joint modeling task that leverages their emotional attributes through multitask learning.

388, **TITLE:** Self Promotion in US Congressional Tweets

<https://www.aclweb.org/anthology/2021.naacl-main.388>

**AUTHORS:** Jun Wang, Kelly Cui, Bei Yu

**HIGHLIGHT:** In this study we built a BERT-based NLP model to predict whether a Congressional tweet shows self-promotion or not and then used this model to examine whether a gender gap in self-promotion exists among Congressional tweets.



- 389, TITLE: Profiling of Intertextuality in Latin Literature Using Word Embeddings  
<https://www.aclweb.org/anthology/2021.naacl-main.389>  
AUTHORS: Patrick J. Burns, James Brofos, Kyle Li, Pramit Chaudhuri, Joseph P. Dexter  
HIGHLIGHT: We report an empirical analysis of intertextuality in classical Latin literature using word embedding models.
- 390, TITLE: Identifying inherent disagreement in natural language inference  
<https://www.aclweb.org/anthology/2021.naacl-main.390>  
AUTHORS: Xinliang Frederick Zhang, Marie-Catherine de Marneffe  
HIGHLIGHT: In this paper, we investigate how to tease systematic inferences (i.e., items for which people agree on the NLI label) apart from disagreement items (i.e., items which lead to different annotations), which most prior work has overlooked.
- 391, TITLE: Modeling Human Mental States with an Entity-based Narrative Graph  
<https://www.aclweb.org/anthology/2021.naacl-main.391>  
AUTHORS: I-Ta Lee, Maria Leonor Pacheco, Dan Goldwasser  
HIGHLIGHT: This paper proposes an Entity-based Narrative Graph (ENG) to model the internal- states of characters in a story.
- 392, TITLE: A Simple and Efficient Multi-Task Learning Approach for Conditioned Dialogue Generation  
<https://www.aclweb.org/anthology/2021.naacl-main.392>  
AUTHORS: Yan Zeng, Jian-Yun Nie  
HIGHLIGHT: In this work, we exploit labeled non-dialogue text data related to the condition, which are much easier to collect.
- 393, TITLE: Hurdles to Progress in Long-form Question Answering  
<https://www.aclweb.org/anthology/2021.naacl-main.393>  
AUTHORS: Kalpesh Krishna, Aurko Roy, Mohit Iyyer  
HIGHLIGHT: While many models have recently been proposed for LFQA, we show in this paper that the task formulation raises fundamental challenges regarding evaluation and dataset creation that currently preclude meaningful modeling progress.
- 394, TITLE: ENTRUST: Argument Reframing with Language Models and Entailment  
<https://www.aclweb.org/anthology/2021.naacl-main.394>  
AUTHORS: Tuhin Chakrabarty, Christopher Hidey, Smaranda Muresan  
HIGHLIGHT: Differences in lexical framing, the focus of our work, can have large effects on peoples' opinions and beliefs. To make progress towards reframing arguments for positive effects, we create a dataset and method for this task.
- 395, TITLE: Paragraph-level Simplification of Medical Texts  
<https://www.aclweb.org/anthology/2021.naacl-main.395>  
AUTHORS: Ashwin Devaraj, Iain Marshall, Byron Wallace, Junyi Jessy Li  
HIGHLIGHT: In this work we introduce a new corpus of parallel texts in English comprising technical and lay summaries of all published evidence pertaining to different clinical topics.
- 396, TITLE: An Empirical Study on Neural Keyphrase Generation  
<https://www.aclweb.org/anthology/2021.naacl-main.396>  
AUTHORS: Rui Meng, Xingdi Yuan, Tong Wang, Sanqiang Zhao, Adam Trischler, Daqing He  
HIGHLIGHT: In this empirical study, we aim to fill this gap by providing extensive experimental results and analyzing the most crucial factors impacting the generalizability of KPG models.
- 397, TITLE: Attention Head Masking for Inference Time Content Selection in Abstractive Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.397>  
AUTHORS: Shuyang Cao, Lu Wang  
HIGHLIGHT: In this work, we present a simple-yet-effective attention head masking technique, which is applied on encoder-decoder attentions to pinpoint salient content at inference time.
- 398, TITLE: Factual Probing Is [MASK]: Learning vs. Learning to Recall  
<https://www.aclweb.org/anthology/2021.naacl-main.398>  
AUTHORS: Zexuan Zhong, Dan Friedman, Danqi Chen  
HIGHLIGHT: In this work, we make two complementary contributions to better understand these factual probing techniques.

- 399, TITLE: Evaluating Saliency Methods for Neural Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.399>  
AUTHORS: Shuoyang Ding, Philipp Koehn  
HIGHLIGHT: To address this question, we conduct a comprehensive and quantitative evaluation of saliency methods on a fundamental category of NLP models: neural language models.
- 400, TITLE: Contextualized Perturbation for Textual Adversarial Attack  
<https://www.aclweb.org/anthology/2021.naacl-main.400>  
AUTHORS: Dianqi Li, Yizhe Zhang, Hao Peng, Liqun Chen, Chris Brockett, Ming-Ting Sun, Bill Dolan  
HIGHLIGHT: This paper presents CLARE, a ContextuaLized AdversaRial Example generation model that produces fluent and grammatical outputs through a mask-then-infill procedure.
- 401, TITLE: DirectProbe: Studying Representations without Classifiers  
<https://www.aclweb.org/anthology/2021.naacl-main.401>  
AUTHORS: Yichu Zhou, Vivek Srikumar  
HIGHLIGHT: In this work, we argue that doing so can be unreliable because different representations may need different classifiers.
- 402, TITLE: Evaluating the Values of Sources in Transfer Learning  
<https://www.aclweb.org/anthology/2021.naacl-main.402>  
AUTHORS: Md Rizwan Parvez, Kai-Wei Chang  
HIGHLIGHT: In this paper, we develop , an efficient source valuation framework for quantifying the usefulness of the sources (e.g., ) in transfer learning based on the Shapley value method.
- 403, TITLE: Too Much in Common: Shifting of Embeddings in Transformer Language Models and its Implications  
<https://www.aclweb.org/anthology/2021.naacl-main.403>  
AUTHORS: Daniel Bis, Maksim Podkorytov, Xiuwen Liu  
HIGHLIGHT: We resolve this by showing, contrary to previous studies, that the representations do not occupy a narrow cone, but rather drift in common directions.
- 404, TITLE: On the Inductive Bias of Masked Language Modeling: From Statistical to Syntactic Dependencies  
<https://www.aclweb.org/anthology/2021.naacl-main.404>  
AUTHORS: Tianyi Zhang, Tatsunori Hashimoto  
HIGHLIGHT: To explain the empirical success of these generic masks, we demonstrate a correspondence between the Masked Language Model (MLM) objective and existing methods for learning statistical dependencies in graphical models.
- 405, TITLE: Limitations of Autoregressive Models and Their Alternatives  
<https://www.aclweb.org/anthology/2021.naacl-main.405>  
AUTHORS: Chu-Cheng Lin, Aaron Jaech, Xin Li, Matthew R. Gormley, Jason Eisner  
HIGHLIGHT: Thus, simply training larger autoregressive language models is not a panacea for NLP. Alternatives include energy-based models (which give up efficient sampling) and latent-variable autoregressive models (which give up efficient scoring of a given string).
- 406, TITLE: On the Transformer Growth for Progressive BERT Training  
<https://www.aclweb.org/anthology/2021.naacl-main.406>  
AUTHORS: Xiaotao Gu, Liyuan Liu, Hongkun Yu, Jing Li, Chen Chen, Jiawei Han  
HIGHLIGHT: Our objective is to help advance the understanding of such Transformer growth and discover principles that guide progressive training.
- 407, TITLE: Revisiting Simple Neural Probabilistic Language Models  
<https://www.aclweb.org/anthology/2021.naacl-main.407>  
AUTHORS: Simeng Sun, Mohit Iyyer  
HIGHLIGHT: In this paper, we revisit the neural probabilistic language model (NPLM) of Bengio et al. (2003), which simply concatenates word embeddings within a fixed window and passes the result through a feed-forward network to predict the next word.
- 408, TITLE: ReadTwice: Reading Very Large Documents with Memories  
<https://www.aclweb.org/anthology/2021.naacl-main.408>  
AUTHORS: Yury Zemlyanskiy, Joshua Ainslie, Michiel de Jong, Philip Pham, Ilya Eckstein, Fei Sha  
HIGHLIGHT: We propose ReadTwice, a simple and effective technique that combines several strengths of prior approaches to model long-range dependencies with Transformers.

- 409, TITLE: SCRIPT: Self-Critic PreTraining of Transformers  
<https://www.aclweb.org/anthology/2021.naacl-main.409>  
AUTHORS: Erik Nijkamp, Bo Pang, Ying Nian Wu, Caiming Xiong  
HIGHLIGHT: We introduce Self-CRItic Pretraining Transformers (SCRIPT) for representation learning of text.
- 410, TITLE: Learning How to Ask: Querying LMs with Mixtures of Soft Prompts  
<https://www.aclweb.org/anthology/2021.naacl-main.410>  
AUTHORS: Guanghui Qin, Jason Eisner  
HIGHLIGHT: We explore the idea of learning prompts by gradient descent-either fine-tuning prompts taken from previous work, or starting from random initialization.
- 411, TITLE: Nutri-bullets Hybrid: Consensual Multi-document Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.411>  
AUTHORS: Darsh Shah, Lili Yu, Tao Lei, Regina Barzilay  
HIGHLIGHT: We present a method for generating comparative summaries that highlight similarities and contradictions in input documents.
- 412, TITLE: AVA: an Automatic eValuation Approach for Question Answering Systems  
<https://www.aclweb.org/anthology/2021.naacl-main.412>  
AUTHORS: Thuy Vu, Alessandro Moschitti  
HIGHLIGHT: We introduce AVA, an automatic evaluation approach for Question Answering, which given a set of questions associated with Gold Standard answers (references), can estimate system Accuracy.
- 413, TITLE: SpanPredict: Extraction of Predictive Document Spans with Neural Attention  
<https://www.aclweb.org/anthology/2021.naacl-main.413>  
AUTHORS: Vivek Subramanian, Matthew Engelhard, Sam Berchuck, Liqun Chen, Ricardo Henao, Lawrence Carin  
HIGHLIGHT: We here formalize this problem as predictive extraction and address it using a simple mechanism based on linear attention.
- 414, TITLE: Text Editing by Command  
<https://www.aclweb.org/anthology/2021.naacl-main.414>  
AUTHORS: Felix Faltings, Michel Galley, Gerold Hintz, Chris Brockett, Chris Quirk, Jianfeng Gao, Bill Dolan  
HIGHLIGHT: To this end, we propose a novel text editing task, and introduce WikiDocEdits, a dataset of single-sentence edits crawled from Wikipedia.
- 415, TITLE: A Deep Metric Learning Approach to Account Linking  
<https://www.aclweb.org/anthology/2021.naacl-main.415>  
AUTHORS: Aleem Khan, Elizabeth Fleming, Noah Schofield, Marcus Bishop, Nicholas Andrews  
HIGHLIGHT: We consider the task of linking social media accounts that belong to the same author in an automated fashion on the basis of the content and meta-data of the corresponding document streams.
- 416, TITLE: Improving Factual Completeness and Consistency of Image-to-Text Radiology Report Generation  
<https://www.aclweb.org/anthology/2021.naacl-main.416>  
AUTHORS: Yasuhide Miura, Yuhao Zhang, Emily Tsai, Curtis Langlotz, Dan Jurafsky  
HIGHLIGHT: Here we introduce two new simple rewards to encourage the generation of factually complete and consistent radiology reports: one that encourages the system to generate radiology domain entities consistent with the reference, and one that uses natural language inference to encourage these entities to be described in inferentially consistent ways.
- 417, TITLE: Multimodal End-to-End Sparse Model for Emotion Recognition  
<https://www.aclweb.org/anthology/2021.naacl-main.417>  
AUTHORS: Wenliang Dai, Samuel Cahyawijaya, Zihan Liu, Pascale Fung  
HIGHLIGHT: In this paper, we develop a fully end-to-end model that connects the two phases and optimizes them jointly.
- 418, TITLE: MIMOQA: Multimodal Input Multimodal Output Question Answering  
<https://www.aclweb.org/anthology/2021.naacl-main.418>  
AUTHORS: Hrituraj Singh, Anshul Nasery, Denil Mehta, Aishwarya Agarwal, Jatin Lamba, Balaji Vasani Srinivasan  
HIGHLIGHT: In this paper, we propose a novel task - MIMOQA - Multimodal Input Multimodal Output Question Answering in which the output is also multimodal.

- 419, TITLE: OCID-Ref: A 3D Robotic Dataset With Embodied Language For Clutter Scene Grounding  
<https://www.aclweb.org/anthology/2021.naacl-main.419>  
AUTHORS: Ke-Jyun Wang, Yun-Hsuan Liu, Hung-Ting Su, Jen-Wei Wang, Yu-Siang Wang, Winston Hsu, Wen-Chin Chen  
HIGHLIGHT: In our work, we propose a novel OCID-Ref dataset featuring a referring expression segmentation task with referring expressions of occluded objects.
- 420, TITLE: Unsupervised Vision-and-Language Pre-training Without Parallel Images and Captions  
<https://www.aclweb.org/anthology/2021.naacl-main.420>  
AUTHORS: Liunian Harold Li, Haoxuan You, Zhecan Wang, Alireza Zareian, Shih-Fu Chang, Kai-Wei Chang  
HIGHLIGHT: In particular, we propose to conduct "mask-and-predict" pre-training on text-only and image-only corpora and introduce the object tags detected by an object recognition model as anchor points to bridge two modalities.
- 421, TITLE: Multitasking Inhibits Semantic Drift  
<https://www.aclweb.org/anthology/2021.naacl-main.421>  
AUTHORS: Athul Paul Jacob, Mike Lewis, Jacob Andreas  
HIGHLIGHT: We study the dynamics of learning in latent language policies (LLPs), in which instructor agents generate natural-language subgoal descriptions and executor agents map these descriptions to low-level actions.
- 422, TITLE: Probing Contextual Language Models for Common Ground with Visual Representations  
<https://www.aclweb.org/anthology/2021.naacl-main.422>  
AUTHORS: Gabriel Ilharco, Rowan Zellers, Ali Farhadi, Hannaneh Hajishirzi  
HIGHLIGHT: In this work, we consider a new question: to what extent contextual representations of concrete nouns are aligned with corresponding visual representations?
- 423, TITLE: BBAEG: Towards BERT-based Biomedical Adversarial Example Generation for Text Classification  
<https://www.aclweb.org/anthology/2021.naacl-main.423>  
AUTHORS: Ishani Mondal  
HIGHLIGHT: We propose BBAEG (Biomedical BERT-based Adversarial Example Generation), a black-box attack algorithm for biomedical text classification, leveraging the strengths of both domain-specific synonym replacement for biomedical named entities and BERT-MLM predictions, spelling variation and number replacement.
- 424, TITLE: Targeted Adversarial Training for Natural Language Understanding  
<https://www.aclweb.org/anthology/2021.naacl-main.424>  
AUTHORS: Lis Pereira, Xiaodong Liu, Hao Cheng, Hoifung Poon, Jianfeng Gao, Ichiro Kobayashi  
HIGHLIGHT: We present a simple yet effective Targeted Adversarial Training (TAT) algorithm to improve adversarial training for natural language understanding.
- 425, TITLE: Latent-Optimized Adversarial Neural Transfer for Sarcasm Detection  
<https://www.aclweb.org/anthology/2021.naacl-main.425>  
AUTHORS: Xu Guo, Boyang Li, Han Yu, Chunyan Miao  
HIGHLIGHT: We propose a generalized latent optimization strategy that allows different losses to accommodate each other and improves training dynamics.
- 426, TITLE: Self-training Improves Pre-training for Natural Language Understanding  
<https://www.aclweb.org/anthology/2021.naacl-main.426>  
AUTHORS: Jingfei Du, Edouard Grave, Beliz Gunel, Vishrav Chaudhary, Onur Celebi, Michael Auli, Veselin Stoyanov, Alexis Conneau  
HIGHLIGHT: In this paper, we study self-training as another way to leverage unlabeled data through semi-supervised learning.
- 427, TITLE: Supporting Clustering with Contrastive Learning  
<https://www.aclweb.org/anthology/2021.naacl-main.427>  
AUTHORS: Dejiao Zhang, Feng Nan, Xiaokai Wei, Shang-Wen Li, Henghui Zhu, Kathleen McKeown, Ramesh Nallapati, Andrew O. Arnold, Bing Xiang  
HIGHLIGHT: To this end, we propose Supporting Clustering with Contrastive Learning (SCCL) - a novel framework to leverage contrastive learning to promote better separation.
- 428, TITLE: TITA: A Two-stage Interaction and Topic-Aware Text Matching Model  
<https://www.aclweb.org/anthology/2021.naacl-main.428>  
AUTHORS: Xingwu Sun, Yanling Cui, Hongyin Tang, Qiuyu Zhu, Fuzheng Zhang, Beihong Jin

**HIGHLIGHT:** In this paper, we focus on the problem of keyword and document matching by considering different relevance levels.

429, **TITLE:** Neural Quality Estimation with Multiple Hypotheses for Grammatical Error Correction

<https://www.aclweb.org/anthology/2021.naacl-main.429>

**AUTHORS:** Zhenghao Liu, Xiaoyuan Yi, Maosong Sun, Liner Yang, Tat-Seng Chua

**HIGHLIGHT:** This paper presents the Neural Verification Network (VERNet) for GEC quality estimation with multiple hypotheses.

430, **TITLE:** Neural Network Surgery: Injecting Data Patterns into Pre-trained Models with Minimal Instance-wise Side Effects

<https://www.aclweb.org/anthology/2021.naacl-main.430>

**AUTHORS:** Zhiyuan Zhang, Xuancheng Ren, Qi Su, Xu Sun, Bin He

**HIGHLIGHT:** Motivated by neuroscientific evidence and theoretical results, we demonstrate that side effects can be controlled by the number of changed parameters and thus, we propose to conduct neural network surgery by only modifying a limited number of parameters.

431, **TITLE:** Discrete Argument Representation Learning for Interactive Argument Pair Identification

<https://www.aclweb.org/anthology/2021.naacl-main.431>

**AUTHORS:** Lu Ji, Zhongyu Wei, Jing Li, Qi Zhang, Xuanjing Huang

**HIGHLIGHT:** In this paper, we focus on identifying interactive argument pairs from two posts with opposite stances to a certain topic.

432, **TITLE:** On Unifying Misinformation Detection

<https://www.aclweb.org/anthology/2021.naacl-main.432>

**AUTHORS:** Nayeon Lee, Belinda Z. Li, Sinong Wang, Pascale Fung, Hao Ma, Wen-tau Yih, Madian Khabsa

**HIGHLIGHT:** In this paper, we introduce UnifiedM2, a general-purpose misinformation model that jointly models multiple domains of misinformation with a single, unified setup.

433, **TITLE:** Frustratingly Easy Edit-based Linguistic Steganography with a Masked Language Model

<https://www.aclweb.org/anthology/2021.naacl-main.433>

**AUTHORS:** Honai Ueoka, Yugo Murawaki, Sadao Kurohashi

**HIGHLIGHT:** In this paper, we revisit edit-based linguistic steganography, with the idea that a masked language model offers an off-the-shelf solution.

434, **TITLE:** Few-Shot Text Classification with Triplet Networks, Data Augmentation, and Curriculum Learning

<https://www.aclweb.org/anthology/2021.naacl-main.434>

**AUTHORS:** Jason Wei, Chengyu Huang, Soroush Vosoughi, Yu Cheng, Shiqi Xu

**HIGHLIGHT:** This paper explores data augmentation-a technique particularly suitable for training with limited data-for this few-shot, highly-multiclass text classification setting.

435, **TITLE:** Do RNN States Encode Abstract Phonological Alternations?

<https://www.aclweb.org/anthology/2021.naacl-main.435>

**AUTHORS:** Miikka Silfverberg, Francis Tyers, Garrett Nicolai, Mans Hulden

**HIGHLIGHT:** To investigate whether complex alternations are simply memorized or whether there is some level of generalization across related sound changes in a sequence-to-sequence model, we perform several experiments on Finnish consonant gradation-a complex set of sound changes triggered in some words by certain suffixes.

436, **TITLE:** Pre-training with Meta Learning for Chinese Word Segmentation

<https://www.aclweb.org/anthology/2021.naacl-main.436>

**AUTHORS:** Zhen Ke, Liang Shi, Songtao Sun, Erli Meng, Bin Wang, Xipeng Qiu

**HIGHLIGHT:** In this paper, we propose a CWS-specific pre-trained model MetaSeg, which employs a unified architecture and incorporates meta learning algorithm into a multi-criteria pre-training task.

437, **TITLE:** Decompose, Fuse and Generate: A Formation-Informed Method for Chinese Definition Generation

<https://www.aclweb.org/anthology/2021.naacl-main.437>

**AUTHORS:** Hua Zheng, Damai Dai, Lei Li, Tianyu Liu, Zhifang Sui, Baobao Chang, Yang Liu

**HIGHLIGHT:** In this paper, we tackle the task of Definition Generation (DG) in Chinese, which aims at automatically generating a definition for a word.

- 438, TITLE: User-Generated Text Corpus for Evaluating Japanese Morphological Analysis and Lexical Normalization  
<https://www.aclweb.org/anthology/2021.naacl-main.438>  
AUTHORS: Shohei Higashiyama, Masao Utiyama, Taro Watanabe, Eiichiro Sumita  
HIGHLIGHT: To evaluate and compare different MA/LN systems, we have constructed a publicly available Japanese UGT corpus.
- 439, TITLE: GPT Perdetry Test: Generating new meanings for new words  
<https://www.aclweb.org/anthology/2021.naacl-main.439>  
AUTHORS: Nikolay Malkin, Sameera Lanka, Pranav Goel, Sudha Rao, Nebojsa Jojic  
HIGHLIGHT: We create a set of nonce words and prompt GPT-3 to generate their dictionary definitions. We find GPT-3 produces plausible definitions that align with human judgments.
- 440, TITLE: Universal Semantic Tagging for English and Mandarin Chinese  
<https://www.aclweb.org/anthology/2021.naacl-main.440>  
AUTHORS: Wenxi Li, Yiyang Hou, Yajie Ye, Li Liang, Weiwei Sun  
HIGHLIGHT: We discuss a set of language-specific semantic phenomena, propose new annotation specifications and build a richly annotated corpus.
- 441, TITLE: ShadowGNN: Graph Projection Neural Network for Text-to-SQL Parser  
<https://www.aclweb.org/anthology/2021.naacl-main.441>  
AUTHORS: Zhi Chen, Lu Chen, Yanbin Zhao, Ruisheng Cao, Zihan Xu, Su Zhu, Kai Yu  
HIGHLIGHT: To improve the model generalization capability for rare and unseen schemas, we propose a new architecture, ShadowGNN, which processes schemas at abstract and semantic levels.
- 442, TITLE: Contextualized and Generalized Sentence Representations by Contrastive Self-Supervised Learning: A Case Study on Discourse Relation Analysis  
<https://www.aclweb.org/anthology/2021.naacl-main.442>  
AUTHORS: Hirokazu Kiyomaru, Sadao Kurohashi  
HIGHLIGHT: We propose a method to learn contextualized and generalized sentence representations using contrastive self-supervised learning.
- 443, TITLE: AMR Parsing with Action-Pointer Transformer  
<https://www.aclweb.org/anthology/2021.naacl-main.443>  
AUTHORS: Jiawei Zhou, Tahira Naseem, Ramon Fernandez Astudillo, Radu Florian  
HIGHLIGHT: In this work, we propose a transition-based system that combines hard-attention over sentences with a target-side action pointer mechanism to decouple source tokens from node representations and address alignments.
- 444, TITLE: NL-EDIT: Correcting Semantic Parse Errors through Natural Language Interaction  
<https://www.aclweb.org/anthology/2021.naacl-main.444>  
AUTHORS: Ahmed Elgohary, Christopher Meek, Matthew Richardson, Adam Fourney, Gonzalo Ramos, Ahmed Hassan Awadallah  
HIGHLIGHT: We present NL-EDIT, a model for interpreting natural language feedback in the interaction context to generate a sequence of edits that can be applied to the initial parse to correct its errors.
- 445, TITLE: Unsupervised Concept Representation Learning for Length-Varying Text Similarity  
<https://www.aclweb.org/anthology/2021.naacl-main.445>  
AUTHORS: Xuchao Zhang, Bo Zong, Wei Cheng, Jingchao Ni, Yanchi Liu, Haifeng Chen  
HIGHLIGHT: In this paper, we propose an unsupervised concept representation learning approach to address the above issues.
- 446, TITLE: Augmenting Knowledge-grounded Conversations with Sequential Knowledge Transition  
<https://www.aclweb.org/anthology/2021.naacl-main.446>  
AUTHORS: Haolan Zhan, Hainan Zhang, Hongshen Chen, Zhuoye Ding, Yongjun Bao, Yanyan Lan  
HIGHLIGHT: To address these problems, we proposed to explicitly model the knowledge transition in sequential multi-turn conversations by abstracting knowledge into topic tags.
- 447, TITLE: Adversarial Self-Supervised Learning for Out-of-Domain Detection  
<https://www.aclweb.org/anthology/2021.naacl-main.447>  
AUTHORS: Zhiyuan Zeng, Keqing He, Yuanmeng Yan, Hong Xu, Weiran Xu  
HIGHLIGHT: To combine the benefits of both types, we propose a self-supervised contrastive learning framework to model discriminative semantic features of both in-domain intents and OOD intents from unlabeled data.

- 448, TITLE: Leveraging Slot Descriptions for Zero-Shot Cross-Domain Dialogue State Tracking  
<https://www.aclweb.org/anthology/2021.naacl-main.448>  
AUTHORS: Zhaojiang Lin, Bing Liu, Seungwhan Moon, Paul Crook, Zhenpeng Zhou, Zhiguang Wang, Zhou Yu, Andrea Madotto, Eunjoon Cho, Rajen Subba  
HIGHLIGHT: In this paper, we propose a slot descriptions enhanced generative approach for zero-shot cross-domain DST.
- 449, TITLE: Hierarchical Transformer for Task Oriented Dialog Systems  
<https://www.aclweb.org/anthology/2021.naacl-main.449>  
AUTHORS: Bishal Santra, Potnuru Anusha, Pawan Goyal  
HIGHLIGHT: In this paper, we propose a generalized framework for Hierarchical Transformer Encoders and show how a standard transformer can be morphed into any hierarchical encoder, including HRED and HIBERT like models, by using specially designed attention masks and positional encodings.
- 450, TITLE: Measuring the 'I don't know' Problem through the Lens of Gricean Quantity  
<https://www.aclweb.org/anthology/2021.naacl-main.450>  
AUTHORS: Huda Khayrallah, Jo?o Sedoc  
HIGHLIGHT: Based on the maxim of Quantity (be informative), we propose Relative Utterance Quantity (RUQ) to diagnose the 'I don't know' problem, in which a dialog system produces generic responses.
- 451, TITLE: RTFE: A Recursive Temporal Fact Embedding Framework for Temporal Knowledge Graph Completion  
<https://www.aclweb.org/anthology/2021.naacl-main.451>  
AUTHORS: Youri Xu, Haihong E, Meina Song, Wenyu Song, Xiaodong Lv, Wang Haotian, Yang Jinrui  
HIGHLIGHT: In this paper, we propose a Recursive Temporal Fact Embedding (RTFE) framework to transplant SKGE models to TKGs and to enhance the performance of existing TKGE models for TKG completion.
- 452, TITLE: Open Hierarchical Relation Extraction  
<https://www.aclweb.org/anthology/2021.naacl-main.452>  
AUTHORS: Kai Zhang, Yuan Yao, Ruobing Xie, Xu Han, Zhiyuan Liu, Fen Lin, Leyu Lin, Maosong Sun  
HIGHLIGHT: To establish the bidirectional connections between OpenRE and relation hierarchy, we propose the task of open hierarchical relation extraction and present a novel OHRE framework for the task.
- 453, TITLE: Jointly Extracting Explicit and Implicit Relational Triples with Reasoning Pattern Enhanced Binary Pointer Network  
<https://www.aclweb.org/anthology/2021.naacl-main.453>  
AUTHORS: Yubo Chen, Yunqi Zhang, Changran Hu, Yongfeng Huang  
HIGHLIGHT: In this paper, we propose a unified framework to jointly extract explicit and implicit relational triples.
- 454, TITLE: Multi-Grained Knowledge Distillation for Named Entity Recognition  
<https://www.aclweb.org/anthology/2021.naacl-main.454>  
AUTHORS: Xuan Zhou, Xiao Zhang, Chenyang Tao, Junya Chen, Bing Xu, Wei Wang, Jing Xiao  
HIGHLIGHT: Drawing power from the recent advance in knowledge distillation (KD), this work presents a novel distillation scheme to efficiently transfer the knowledge learned from big models to their more affordable counterpart.
- 455, TITLE: SGG: Learning to Select, Guide, and Generate for Keyphrase Generation  
<https://www.aclweb.org/anthology/2021.naacl-main.455>  
AUTHORS: Jing Zhao, Junwei Bao, Yifan Wang, Youzheng Wu, Xiaodong He, Bowen Zhou  
HIGHLIGHT: In this paper, a Select-Guide-Generate (SGG) approach is proposed to deal with present and absent keyphrases generation separately with different mechanisms.
- 456, TITLE: Towards Sentiment and Emotion aided Multi-modal Speech Act Classification in Twitter  
<https://www.aclweb.org/anthology/2021.naacl-main.456>  
AUTHORS: Tulika Saha, Apoorva Upadhyaya, Sriparna Saha, Pushpak Bhattacharyya  
HIGHLIGHT: We propose a Dyadic Attention Mechanism (DAM) based multi-modal, adversarial multi-tasking framework.
- 457, TITLE: Generative Imagination Elevates Machine Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.457>  
AUTHORS: Quanyu Long, Mingxuan Wang, Lei Li  
HIGHLIGHT: In this paper, we propose ImagiT, a novel machine translation method via visual imagination.

- 458, TITLE: Non-Autoregressive Translation by Learning Target Categorical Codes  
<https://www.aclweb.org/anthology/2021.naacl-main.458>  
AUTHORS: Yu Bao, Shujian Huang, Tong Xiao, Dongqi Wang, Xinyu Dai, Jiajun Chen  
HIGHLIGHT: In this paper, we propose CNAT, which learns implicitly categorical codes as latent variables into the non-autoregressive decoding.
- 459, TITLE: Training Data Augmentation for Code-Mixed Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.459>  
AUTHORS: Abhirut Gupta, Aditya Vavre, Sunita Sarawagi  
HIGHLIGHT: We present an m-BERT based procedure whose core learnable component is a ternary sequence labeling model, that can be trained with a limited code-mixed corpus alone.
- 460, TITLE: Rethinking Perturbations in Encoder-Decoders for Fast Training  
<https://www.aclweb.org/anthology/2021.naacl-main.460>  
AUTHORS: Sho Takase, Shun Kiyono  
HIGHLIGHT: Thus, this study addresses the question of whether these approaches are efficient enough for training time.
- 461, TITLE: Context-aware Decoder for Neural Machine Translation using a Target-side Document-Level Language Model  
<https://www.aclweb.org/anthology/2021.naacl-main.461>  
AUTHORS: Amane Sugiyama, Naoki Yoshinaga  
HIGHLIGHT: We therefore present a simple method to perform context-aware decoding with any pre-trained sentence-level translation model by using a document-level language model.
- 462, TITLE: Machine Translated Text Detection Through Text Similarity with Round-Trip Translation  
<https://www.aclweb.org/anthology/2021.naacl-main.462>  
AUTHORS: Hoang-Quoc Nguyen-Son, Tran Thao, Seira Hidano, Ishita Gupta, Shinsaku Kiyomoto  
HIGHLIGHT: Hence, we propose a detector using text similarity with round-trip translation (TSRT).
- 463, TITLE: TR-BERT: Dynamic Token Reduction for Accelerating BERT Inference  
<https://www.aclweb.org/anthology/2021.naacl-main.463>  
AUTHORS: Deming Ye, Yankai Lin, Yufei Huang, Maosong Sun  
HIGHLIGHT: To address this issue, we propose a dynamic token reduction approach to accelerate PLMs' inference, named TR-BERT, which could flexibly adapt the layer number of each token in inference to avoid redundant calculation.
- 464, TITLE: Breadth First Reasoning Graph for Multi-hop Question Answering  
<https://www.aclweb.org/anthology/2021.naacl-main.464>  
AUTHORS: Yongjie Huang, Meng Yang  
HIGHLIGHT: In this paper, we propose a novel model of Breadth First Reasoning Graph (BFR-Graph), which presents a new message passing way that better conforms to the reasoning process.
- 465, TITLE: Improving Zero-Shot Cross-lingual Transfer for Multilingual Question Answering over Knowledge Graph  
<https://www.aclweb.org/anthology/2021.naacl-main.465>  
AUTHORS: Yucheng Zhou, Xiubo Geng, Tao Shen, Wenqiang Zhang, Daxin Jiang  
HIGHLIGHT: In this paper, we exploit unsupervised bilingual lexicon induction (BLI) to map training questions in source language into those in target language as augmented training data, which circumvents language inconsistency between training and inference.
- 466, TITLE: RocketQA: An Optimized Training Approach to Dense Passage Retrieval for Open-Domain Question Answering  
<https://www.aclweb.org/anthology/2021.naacl-main.466>  
AUTHORS: Yingqi Qu, Yuchen Ding, Jing Liu, Kai Liu, Ruiyang Ren, Wayne Xin Zhao, Daxiang Dong, Hua Wu, Haifeng Wang  
HIGHLIGHT: To address these challenges, we propose an optimized training approach, called RocketQA, to improving dense passage retrieval.
- 467, TITLE: DAGN: Discourse-Aware Graph Network for Logical Reasoning  
<https://www.aclweb.org/anthology/2021.naacl-main.467>  
AUTHORS: Yinya Huang, Meng Fang, Yu Cao, Liwei Wang, Xiaodan Liang  
HIGHLIGHT: In this work, we explore aggregating passage-level clues for solving logical reasoning QA by using discourse-based information.



- 468, TITLE: Designing a Minimal Retrieve-and-Read System for Open-Domain Question Answering  
<https://www.aclweb.org/anthology/2021.naacl-main.468>  
AUTHORS: Sohee Yang, Minjoon Seo  
HIGHLIGHT: Here, we discuss several orthogonal strategies to drastically reduce the footprint of a retrieve-and-read open-domain QA system by up to 160x.
- 469, TITLE: Unsupervised Multi-hop Question Answering by Question Generation  
<https://www.aclweb.org/anthology/2021.naacl-main.469>  
AUTHORS: Liangming Pan, Wenhui Chen, Wenhan Xiong, Min-Yen Kan, William Yang Wang  
HIGHLIGHT: We propose MQA-QG, an unsupervised framework that can generate human-like multi-hop training data from both homogeneous and heterogeneous data sources.
- 470, TITLE: Sliding Selector Network with Dynamic Memory for Extractive Summarization of Long Documents  
<https://www.aclweb.org/anthology/2021.naacl-main.470>  
AUTHORS: Peng Cui, Le Hu  
HIGHLIGHT: To address this issue, we propose the sliding selector network with dynamic memory for extractive summarization of long-form documents, which employs a sliding window to extract summary sentences segment by segment.
- 471, TITLE: AdaptSum: Towards Low-Resource Domain Adaptation for Abstractive Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.471>  
AUTHORS: Tiezheng Yu, Zihan Liu, Pascale Fung  
HIGHLIGHT: In this paper, we present a study of domain adaptation for the abstractive summarization task across six diverse target domains in a low-resource setting.
- 472, TITLE: QMSum: A New Benchmark for Query-based Multi-domain Meeting Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.472>  
AUTHORS: Ming Zhong, Da Yin, Tao Yu, Ahmad Zaidi, Mutethia Mutuma, Rahul Jha, Ahmed Hassan Awadallah, Asli Celikyilmaz, Yang Liu, Xipeng Qiu, Dragomir Radev  
HIGHLIGHT: In order to satisfy the needs of different types of users, we define a new query-based multi-domain meeting summarization task, where models have to select and summarize relevant spans of meetings in response to a query, and we introduce QMSum, a new benchmark for this task.
- 473, TITLE: MM-AVS: A Full-Scale Dataset for Multi-modal Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.473>  
AUTHORS: Xiyan Fu, Jun Wang, Zhenglu Yang  
HIGHLIGHT: In this study, we release a full-scale multimodal dataset comprehensively gathering documents, summaries, images, captions, videos, audios, transcripts, and titles in English from CNN and Daily Mail.
- 474, TITLE: MediaSum: A Large-scale Media Interview Dataset for Dialogue Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.474>  
AUTHORS: Chenguang Zhu, Yang Liu, Jie Mei, Michael Zeng  
HIGHLIGHT: This paper introduces MediaSum, a large-scale media interview dataset consisting of 463.6K transcripts with abstractive summaries.
- 475, TITLE: Improving Faithfulness in Abstractive Summarization with Contrast Candidate Generation and Selection  
<https://www.aclweb.org/anthology/2021.naacl-main.475>  
AUTHORS: Sihao Chen, Fan Zhang, Kazuo Sone, Dan Roth  
HIGHLIGHT: To address the issue, we study contrast candidate generation and selection as a model-agnostic post-processing technique to correct the extrinsic hallucinations (i.e. information not present in the source text) in unfaithful summaries.
- 476, TITLE: Inference Time Style Control for Summarization  
<https://www.aclweb.org/anthology/2021.naacl-main.476>  
AUTHORS: Shuyang Cao, Lu Wang  
HIGHLIGHT: We present two novel methods that can be deployed during summary decoding on any pre-trained Transformer-based summarization model.