

NAACL-HLT 2021

**The 2021 Conference
of the North American Chapter
of the Association for Computational Linguistics:
Human Language Technologies**

Industry Papers

June 6 - 11, 2021

©2021 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-954085-47-3

Message from the Industry Track Chairs

Language technologies and their applications are an integral and critical part of our daily lives. The development of many of these technologies trace their roots to academic and industrial research laboratories where researchers invented a plethora of algorithms, benchmarked them against shared datasets and perfected the performance of these algorithms to provide plausible solutions to real-world applications. While a controlled laboratory setting is vital for a deeper scientific understanding of the language problem and the impact of algorithmic design choices on the performance of a technology, transitioning the technology to real-world industrial strength applications raises a different, yet challenging, set of technical issues.

The Industry Track at NAACL HLT 2021 represents innovations and implementations in speech and natural language processing technologies and systems that are relevant to industrial applications. The primary focus of this track is on papers that advance the understanding of, and demonstrate the effective handling of, practical issues related to the deployment of language processing technologies in non-trivial real-world systems. By “non-trivial real-world system” we mean an application that is deployed for real-world use, i.e. outside controlled environments such as a laboratories, classrooms or experimental crowd-sourced setups, and that uses natural language processing (including speech technology), even if not state of the art in terms of research. There is no requirement that the system be made by a for-profit company, but the users of the system must be outside of the NLP research community.

We received 132 papers, and accepted 39, or 29.5%. This year, nearly half the papers in the Industry Track are about dialog processing, reflecting the importance of interactive systems in industrial applications today. Additional topics include semantics, machine translation, information retrieval, and sentiment analysis.

We hope that these papers will provide an interesting complement to the Main Track papers.

Young-bum Kim, Yunyao Li, Owen Rambow

Organizing Committee

Young-bum Kim, Amazon, Inc. (youngbum@amazon.com)

Yunyaoli Li, IBM, Inc. (yunyaoli@us.ibm.com)

Owen Rambow, Stony Brook University (owen.rambow@stonybrook.edu)

Reviewers

We thank our reviewers, without whom the Industry Track would not be possible. This list also includes reviewers for the secondary ethics review.

Jade Abbott, Mohamed Abdelhady, George Acquaaah-Mensah, Gilles Adda, Shazia Afzal, Sachin Agarwal, Hua Ai, Alan Akbik, Mohamed ALTantawy, Shankar Ananthakrishnan, Ankit Arun, Venkatesh Baglodi, Srinivas Bangalore, Nikoletta Basiou, Daniel Bauer, Frederic Bechet, Tilman Becker, Emily M. Bender, Luciana Benotti, Dan Bikel, Claudia Borg, Trung Bui, Greg Burnham, Donna Byron, Aoife Cahill, Vitor Carvalho, Thiago Castro Ferreira, Sourish Chaudhuri, Ciprian Chelba, John Chen, Minhua Chen, Laura Chiticariu, Justin Chiu, Eunah Cho, Jaesik Choi, Jaegul Choo, Jennifer Chu-Carroll, Hyung Won Chung, Rylan Conway, Deborah Dahl, Marina Danilevsky, Budhaditya Deb, Lingjia Deng, Giuseppe Di Fabbri, Christine Doran, Dejing Dou, Pablo Duboue, Matthew Dunn, David Elson, Ramy Eskander, Xing Fan, Song Feng, Oliver Ferschke, Michael Flor, Karën Fort, Annemarie Friedrich, Ankur Gandhe, Rashmi Gangadharaiah, Judith Gaspers, Anna Lisa Gentile, Ryan Georgi, Debanjan Ghosh, Honglei Guo, Yufan Guo, Daniel Hardt, Hua He, Enrique Henestroza Anguiano, Sanjika Hewavitharana, Christopher Hidey, Derrick Higgins, Lynette Hirschman, Yufang Hou, Samar Husain, Javid Huseynov, Hyesung Ji, Hongxia Jin, Mahesh Joshi, Mohammad Kachuee, Jun Seok Kang, Yoav Katz, Saurabh Khanwalkar, Sun Kim, Dongchan Kim, Yoon Kim, Joo-Kyung Kim, Yu-Seop Kim, Jin-Dong Kim, Jared Kramer, Sanjeev Kumar, Anjishnu Kumar, Rohit Kumar, Gakuto Kurata, Sarasi Lalithsena, Anastassia Lastname, Young-Suk Lee, Han Li, Constantine Lignos, Heuseok Lim, Chin-Yew Lin, Xiaohu Liu, Anastassia Loukina, Alex Marin, Yuval Marton, Yuji Matsumoto, David McDonald, Angeliki Metallinou, Lisa Michaud, Margot Mieskes, Nyalleng Moorosi, Michelle Morales, Isabelle Moulinier, Matthew Mulholland, Udhyakumar Nallasamy, Jinseok Nam, Nobal B. Niraula, Elnaz Nouri, Mari Olsen, Cecile Paris, Youngja Park, Taiwoo Park, Dookun Park, Patrick Paroubek, Ioannis Partalas, Siddharth Patwardhan, Karl Pichotta, Vassilis Plachouras, Alexandros Potamianos, Saloni Potdar, Rashmi Prasad, Long Qin, Elio Querze, Sravana Reddy, Ehud Reiter, Nicholas Ruiz, Alicia Sagae, Avneesh Saluja, Mark Sammons, Ruhi Sarikaya, Hassan Sawaf, Frank Schilder, Ethan Selfridge, Igor Shalyminov, Michal Shmueli-Scheuer, Lei Shu, Sunayana Sitaram, AJ Stent, Svetlana Stoyanchev, Chengwei Su, Tara Taghavi, Joel Tetreault, Sudarshan R. Thitte, Isabel Trancoso, Keith Trnka, Ling Tsou, Morgan Ulinski, Ngoc Phuoc An Vo, Yi-Chia Wang, Lucy Lu Wang, Guoyin Wang, Kyle Williams, Puyang Xu, Yeongyook Yang, Kai Yu, Seunghak Yu, Liyuan Zhang, Yefeng Zheng

Table of Contents

<i>When does text prediction benefit from additional context? An exploration of contextual signals for chat and email messages</i>	
Stojan Trajanovski, Chad Atalla, Kunho Kim, Vipul Agarwal, Milad Shokouhi and Chris Quirk . . .	1
<i>Identifying and Resolving Annotation Changes for Natural Language Understanding</i>	
Jose Garrido Ramas, Giorgio Pessot, Abdalghani Abujabal and Martin Rajman	10
<i>Optimizing NLU Reranking Using Entity Resolution Signals in Multi-domain Dialog Systems</i>	
Tong Wang, Jiangning Chen, Mohsen Malmir, Shuyan Dong, Xin He, Han Wang, Chengwei Su, Yue Liu and Yang Liu	19
<i>Entity Resolution in Open-domain Conversations</i>	
Mingyue Shang, Tong Wang, Mihail Eric, Jiangning Chen, Jiyang Wang, Matthew Welch, Tiantong Deng, Akshay Grewal, Han Wang, Yue Liu, Yang Liu and Dilek Hakkani-Tur	26
<i>Pretrain-Finetune Based Training of Task-Oriented Dialogue Systems in a Real-World Setting</i>	
Manisha Srivastava, Yichao Lu, Riley Peschon and Chenyang Li	34
<i>Contextual Domain Classification with Temporal Representations</i>	
Tzu-Hsiang Lin, Yipeng Shi, Chentao Ye, Yang Fan, Weitong Ruan, Emre Barut, Wael Hamza and Chengwei Su	41
<i>Bootstrapping a Music Voice Assistant with Weak Supervision</i>	
Sergio Oramas, Massimo Quadrana and Fabien Gouyon	49
<i>Continuous Model Improvement for Language Understanding with Machine Translation</i>	
Abdalghani Abujabal, Claudio Delli Bovi, Sungho Ryu, Turan Gojayev, Fabian Triefenbach and Yannick Versley	56
<i>A Hybrid Approach to Scalable and Robust Spoken Language Understanding in Enterprise Virtual Agents</i>	
Ryan Price, Mahnoosh Mehrabani, Narendra Gupta, Yeon-Jun Kim, Shahab Jalalvand, Minhua Chen, Yanjie Zhao and Srinivas Bangalore	63
<i>Protono: Text Normalization with Limited Data for Fast Deployment in Text to Speech Systems</i>	
Shubhi Tyagi, Antonio Bonafonte, Jaime Lorenzo-Trueba and Javier Latorre	72
<i>Addressing the Vulnerability of NMT in Input Perturbations</i>	
Weiwen Xu, Ai Ti Aw, Yang Ding, Kui Wu and Shafiq Joty	80
<i>Cross-lingual Supervision Improves Unsupervised Neural Machine Translation</i>	
Mingxuan Wang, Hongxiao Bai, Lei Li and Hai Zhao	89
<i>Should we find another model?: Improving Neural Machine Translation Performance with ONE-Piece Tokenization Method without Model Modification</i>	
chanjun park, Sugyeong Eo, Hyeonseok Moon and Heuseok Lim	97
<i>Autocorrect in the Process of Translation — Multi-task Learning Improves Dialogue Machine Translation</i>	
Tao Wang, Chengqi Zhao, Mingxuan Wang, Lei Li and Deyi Xiong	105

<i>LightSeq: A High Performance Inference Library for Transformers</i> Xiaohui Wang, Ying Xiong, Yang Wei, Mingxuan Wang and Lei Li	113
<i>Practical Transformer-based Multilingual Text Classification</i> Cindy Wang and Michele Banko	121
<i>An Emotional Comfort Framework for Improving User Satisfaction in E-Commerce Customer Service Chatbots</i> Shuangyong Song, Chao Wang, Haiqing Chen and Huan Chen	130
<i>Language Scaling for Universal Suggested Replies Model</i> Qianlan Ying, Payal Bajaj, Budhaditya Deb, Yu Yang, Wei Wang, Bojia Lin, Milad Shokouhi, Xia Song, Yang Yang and Daxin Jiang	138
<i>Graph-based Multilingual Product Retrieval in E-Commerce Search</i> Hanqing Lu, Youna Hu, Tong Zhao, Tony Wu, Yiwei Song and Bing Yin	146
<i>Query2Prod2Vec: Grounded Word Embeddings for eCommerce</i> Federico Bianchi, Jacopo Tagliabue and Bingqing Yu	154
<i>An Architecture for Accelerated Large-Scale Inference of Transformer-Based Language Models</i> Amir Ganiev, Colton Chapin, Anderson De Andrade and Chen Liu	163
<i>When and Why a Model Fails? A Human-in-the-loop Error Detection Framework for Sentiment Analysis</i> Zhe Liu, Yufan Guo and Jalal Mahmud	170
<i>Technical Question Answering across Tasks and Domains</i> Wenhao Yu, Lingfei Wu, Yu Deng, Qingkai Zeng, Ruchi Mahindru, Sinem Guven and Meng Jiang	178
<i>Cost-effective Deployment of BERT Models in Serverless Environment</i> Marek Suppa, Katarína Benešová and Andrej Švec	187
<i>Noise Robust Named Entity Understanding for Voice Assistants</i> Deepak Muralidharan, Joel Ruben Antony Moniz, Sida Gao, Xiao Yang, Justine Kao, Stephen Pulman, Atish Kothari, Ray Shen, Yinying Pan, Vivek Kaul, Mubarak Seyed Ibrahim, Gang Xiang, Nan Dun, Yidan Zhou, Andy O, Yuan Zhang, Pooja Chitkara, Xuan Wang, Alkesh Patel, Kushal Tayal, Roger Zheng, Peter Gräsch, Jason D Williams and Lin Li	196
<i>Goodwill Hunting: Analyzing and Repurposing Off-the-Shelf Named Entity Linking Systems</i> Karan Goel, Laurel Orr, Nazneen Fatema Rajani, Jesse Vig and Christopher Ré	205
<i>Intent Features for Rich Natural Language Understanding</i> Brian Lester, Sagnik Ray Choudhury, Rashmi Prasad and Srinivas Bangalore	214
<i>Development of an Enterprise-Grade Contract Understanding System</i> Arvind Agarwal, Laura Chiticariu, Poornima Chozhiyath Raman, Marina Danilevsky, Diman Ghazi, Ankush Gupta, Shanmukha Guttula, Yannis Katsis, Rajasekar Krishnamurthy, Yunyao Li, Shubham Mudgal, Vitobha Munigala, Nicholas Phan, Dhaval Sonawane, Sneha Srinivasan, Sudarshan R. Thitte, Mitesh Vasa, Ramiya Venkatachalam, Vinitha Yaski and Huaiyu Zhu	222
<i>Discovering Better Model Architectures for Medical Query Understanding</i> Wei Zhu, Yuan Ni, Xiaoling Wang and Guotong Xie	230

<i>OodGAN: Generative Adversarial Network for Out-of-Domain Data Generation</i> Petr Marek, Vishal Ishwar Naik, Anuj Goyal and Vincent Auvray	238
<i>Coherent and Concise Radiology Report Generation via Context Specific Image Representations and Orthogonal Sentence States</i> Litton J Kurisinkel, Ai Ti Aw and Nancy F Chen	246
<i>An Empirical Study of Generating Texts for Search Engine Advertising</i> Hidetaka Kamigaito, Peinan Zhang, Hiroya Takamura and Manabu Okumura	255
<i>Ad Headline Generation using Self-Critical Masked Language Model</i> Yashal Shakti Kanungo, Sumit Negi and Aruna Rajan	263
<i>LATEX-Numeric: Language Agnostic Text Attribute Extraction for Numeric Attributes</i> Kartik Mehta, Ioana Oprea and Nikhil Rasiwasia	272
<i>Training Language Models under Resource Constraints for Adversarial Advertisement Detection</i> Eshwar Shamanna Girishekar, Shiv Surya, Nishant Nikhil, Dyut Kumar Sil, Sumit Negi and Aruna Rajan	280
<i>Combining Weakly Supervised ML Techniques for Low-Resource NLU</i> Victor Soto and Konstantine Arkoudas	288
<i>Label-Guided Learning for Item Categorization in e-Commerce</i> Lei Chen and Hirokazu Miyake	296
<i>Benchmarking Commercial Intent Detection Services with Practice-Driven Evaluations</i> Haode Qi, Lin Pan, Atin Sood, Abhishek Shah, Ladislav Kunc, Mo Yu and Saloni Potdar	304
<i>Industry Scale Semi-Supervised Learning for Natural Language Understanding</i> Luoxin Chen, Francisco Garcia, Varun Kumar, He Xie and Jianhua Lu	311

Industry Track Program

Mon 07 Jun 2021 (all times PDT, UTC-7)

09:00–10:20 IND1-Oral: Dialogue (Industry Track)

When does text prediction benefit from additional context? An exploration of contextual signals for chat and email messages

Stojan Trajanovski, Chad Atalla, Kunho Kim, Vipul Agarwal, Milad Shokouhi and Chris Quirk

Identifying and Resolving Annotation Changes for Natural Language Understanding

Jose Garrido Ramas, Giorgio Pessot, Abdalghani Abujabal and Martin Rajman

Optimizing NLU Reranking Using Entity Resolution Signals in Multi-domain Dialog Systems

Tong Wang, Jiangning Chen, Mohsen Malmir, Shuyan Dong, Xin He, Han Wang, Chengwei Su, Yue Liu and Yang Liu

Entity Resolution in Open-domain Conversations

Mingyue Shang, Tong Wang, Mihail Eric, Jiangning Chen, Jiyang Wang, Matthew Welch, Tiantong Deng, Akshay Grewal, Han Wang, Yue Liu, Yang Liu and Dilek Hakkani-Tur

Pretrain-Finetune Based Training of Task-Oriented Dialogue Systems in a Real-World Setting

Manisha Srivastava, Yichao Lu, Riley Peschon and Chenyang Li

Contextual Domain Classification with Temporal Representations

Tzu-Hsiang Lin, Yipeng Shi, Chentao Ye, Yang Fan, Weitong Ruan, Emre Barut, Wael Hamza and Chengwei Su

10:20–11:40 IND2-Oral: Spoken Dialogue (Industry Track)

Bootstrapping a Music Voice Assistant with Weak Supervision

Sergio Oramas, Massimo Quadrana and Fabien Gouyon

Continuous Model Improvement for Language Understanding with Machine Translation

Abdalghani Abujabal, Claudio Delli Bovi, Sungho Ryu, Turan Gojayev, Fabian Triefenbach and Yannick Versley

A HYBRID APPROACH TO SCALABLE AND ROBUST SPOKEN LANGUAGE UNDERSTANDING IN ENTERPRISE VIRTUAL AGENTS

Ryan Price, Mahnoosh Mehrabani, Narendra Gupta, Yeon-Jun Kim, Shahab Jalalvand, Minhua Chen, Yanjie Zhao and Srinivas Bangalore

Proteno: Text Normalization with Limited Data for Fast Deployment in Text to Speech Systems

Shubhi Tyagi, Antonio Bonafonte, Jaime Lorenzo-Trueba and Javier Latorre

Mon 07 Jun 2021 (continued; all times PDT, UTC-7))

18:20–19:40 IND3-Oral: Machine Translation (Industry Track)

Addressing the Vulnerability of NMT in Input Perturbations

Weiwen Xu, AiTi Aw, Yang Ding, Kui Wu and Shafiq Joty

Cross-lingual Supervision Improves Unsupervised Neural Machine Translation

Mingxuan Wang, Hongxiao Bai, Lei Li and Hai Zhao

Should we find another model?: Improving Neural Machine Translation Performance with ONE-Piece Tokenization Method without Model Modification

chanjun park, Sugyeong Eo, Hyeonseok Moon and Heuseok Lim

Autocorrect in the Process of Translation — Multi-task Learning Improves Dialogue Machine Translation

Tao Wang, Chengqi Zhao, Mingxuan Wang, Lei Li and Deyi Xiong

LightSeq: A High Performance Inference Library for Transformers

Xiaohui Wang, Ying Xiong, Yang Wei, Mingxuan Wang and Lei Li

Practical Transformer-based Multilingual Text Classification

Cindy Wang and Michele Banko

19:40–21:00 IND4-Oral: Sentiment Analysis and Information Retrieval (Industry Track)

An Emotional Comfort Framework for Improving User Satisfaction in E-Commerce Customer Service Chatbots

Shuangyong Song, Chao Wang, Haiqing Chen and Huan Chen

Language Scaling for Universal Suggested Replies Model

Qianlan Ying, Payal Bajaj, Budhadya Deb, Yu Yang, Wei Wang, Bojia Lin, Milad Shokouhi, Xia Song, Yang Yang and Daxin Jiang

Graph-based Multilingual Product Retrieval in E-Commerce Search

Hanqing Lu, Youna Hu, Tong Zhao, Tony Wu, Yiwei Song and Bing Yin

Query2Prod2Vec: Grounded Word Embeddings for eCommerce

Federico Bianchi, Jacopo Tagliabue and Bingqing Yu

Mon 07 Jun 2021 (continued; all times PDT, UTC-7)

An Architecture for Accelerated Large-Scale Inference of Transformer-Based Language Models

Amir Ganiev, Colton Chapin, Anderson De Andrade and Chen Liu

When and Why a Model Fails? A Human-in-the-loop Error Detection Framework for Sentiment Analysis

Zhe Liu, Yufan Guo and Jalal Mahmud

Tue 08 Jun 2021 (all times PDT, UTC-7)

09:00–10:20 IND5-Oral: Semantics (Industry Track)

Technical Question Answering across Tasks and Domains

Wenhao Yu, Lingfei Wu, Yu Deng, Qingkai Zeng, Ruchi Mahindru, Sinem Guven and Meng Jiang

Cost-effective Deployment of BERT Models in Serverless Environment

Marek Suppa, Katarína Benešová and Andrej Švec

Noise Robust Named Entity Understanding for Voice Assistants

Deepak Muralidharan, Joel Ruben Antony Moniz, Sida Gao, Xiao Yang, Justine Kao, Stephen Pulman, Atish Kothari, Ray Shen, Yinying Pan, Vivek Kaul, Mubarak Seyed Ibrahim, Gang Xiang, Nan Dun, Yidan Zhou, Andy O, Yuan Zhang, Pooja Chitkara, Xuan Wang, Alkesh Patel, Kushal Tayal, Roger Zheng, Peter Gräsch, Jason D Williams and Lin Li

Goodwill Hunting: Analyzing and Repurposing Off-the-Shelf Named Entity Linking Systems

Karan Goel, Laurel Orr, Nazneen Fatema Rajani, Jesse Vig and Christopher Ré

Intent Features for Rich Natural Language Understanding

Brian Lester, Sagnik Ray Choudhury, Rashmi Prasad and Srinivas Bangalore

Development of an Enterprise-Grade Contract Understanding System

Arvind Agarwal, Laura Chiticariu, Poornima Chozhiyath Raman, Marina Danilevsky, Diman Ghazi, Ankush Gupta, Shanmukha Guttula, Yannis Katsis, Rajasekar Krishnamurthy, Yunyao Li, Shubham Mudgal, Vitobha Munigala, Nicholas Phan, Dhaval Sonawane, Sneha Srinivasan, Sudarshan R. Thitte, Mitesh Vasa, Ramiya Venkatachalam, Vinitha Yaski and Huaiyu Zhu

Wed 09 Jun 2021 (all times PDT, UTC-7)

18:20–19:40 IND6-Oral: Natural Language Generation and Dialogue (Industry Track)

Discovering Better Model Architectures for Medical Query Understanding

Wei Zhu, Yuan Ni, Xiaoling Wang and Guotong Xie

OodGAN: Generative Adversarial Network for Out-of-Domain Data Generation

Petr Marek, Vishal Ishwar Naik, Anuj Goyal and Vincent Auvray

Coherent and Concise Radiology Report Generation via Context Specific Image Representations and Orthogonal Sentence States

Litton J Kurisinkel, Ai Ti Aw and Nancy F Chen

An Empirical Study of Generating Texts for Search Engine Advertising

Hidetaka Kamigaito, Peinan Zhang, Hiroya Takamura and Manabu Okumura

Ad Headline Generation using Self-Critical Masked Language Model

Yashal Shakti Kanungo, Sumit Negi and Aruna Rajan

19:40–21:00 IND7-Oral: Machine Learning (Industry Track)

LATEX-Numeric: Language Agnostic Text Attribute Extraction for Numeric Attributes

Kartik Mehta, Ioana Oprea and Nikhil Rasiwasia

Training Language Models under Resource Constraints for Adversarial Advertisement Detection

Eshwar Shamanna Girishekar, Shiv Surya, Nishant Nikhil, Dyut Kumar Sil, Sumit Negi and Aruna Rajan

Combining Weakly Supervised ML Techniques for Low-Resource NLU

Victor Soto and Konstantine Arkoudas

Label-Guided Learning for Item Categorization in e-Commerce

Lei Chen and Hirokazu Miyake

Benchmarking Commercial Intent Detection Services with Practice-Driven Evaluations

Haode Qi, Lin Pan, Atin Sood, Abhishek Shah, Ladislav Kunc, Mo Yu and Saloni Potdar

Wed 09 Jun 2021 (continued; all times PDT, UTC-7)

Industry Scale Semi-Supervised Learning for Natural Language Understanding
Luoxin Chen, Francisco Garcia, Varun Kumar, He Xie and Jianhua Lu