

NAACL HLT 2019

**The 2019 Conference of the
North American Chapter of the
Association for Computational Linguistics:
Human Language Technologies**

**Proceedings of the Conference
Vol. 2 (Industry Papers)**

June 2 - June 7, 2019

©2019 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 978-1-950737-14-7

Introduction

We welcome you to the second installment of the NAACL-HLT Industry Track.

Introduced at NAACL-HLT 2018 in New Orleans, the industry track provides a forum for researchers, engineers and application developers to exchange ideas, share results and discuss use cases of successful deployment of language technologies in real-world settings. The inaugural Industry track was very successful in terms of both participation during the conference and feedback received through the post-conference survey.

Continuing the industry track into its second year, we took stock of the opinions that participants and organizers shared about the separate industry track. Many recognized the need to reflect the shift in contemporary NLP work which has grown substantially outside academic and research organizations into industry. On the other hand, there was a concern that the industry track unfairly advantaged industry-affiliated participants. This year we have taken small steps towards integrating the industry track more into the conference by adhering to the same standards as the main track of the conference in terms of timelines and acceptance ratios. At the same time, we have made conscious attempts to welcome all constituents of the NAACL conferences to this new track.

Submissions to the industry track were solicited from all members of the NAACL community including but not limited to students, practitioners and researchers. The call for papers focused on advances and challenges in the deployment of language processing technologies in real-world systems. Following the same deadlines and policies as the main track of the NAACL-HLT 2019 conference, we have aimed to eliminate the perceptions of unfair advantages to papers submitted to the industry track.

Despite moving the submission deadline to over a month earlier and requiring that the authors choose the most suitable track for their papers, we received 124 abstracts and 114 paper submissions, a 25% year-over-year increase. Twelve percent of the submissions were rejected without review due to incompleteness, non-compliance with format requirements or submission policies (such as the double submission policy). We saw a remarkable increase in papers that were co-authored by researchers in academia and industry labs: 48% vs. 29% in 2018.

Submitted papers were reviewed by our program committee with rich representation of the present spectrum of NLP researchers and professionals. Each submission was reviewed by at least three members of the program committee. Reviews solicited committee opinions along two primary aspects: focus on real-world applications and lessons offered by the paper. Reviews also took into consideration clarity, methodological rigor, ethical use of datasets and compliance with conference guidelines. Thanks to the enthusiastic and diligent efforts of the industry track program committee, the reviews were completed on time. We accepted 28 papers based on committee recommendations as well as alignment of the papers with the goals of industry track. The acceptance rate reduced from 33% to 28% compared to NAACL-HLT 2018.

The Industry Track program this year will consist of two oral sessions (5 papers each) and one poster session (18 posters). The presentation format was determined based on reviewer recommendations as well as a paper's overall score. The first oral session will address various challenges of using language technologies in real-world applications. One of the common themes for many papers in this session is ensuring system robustness towards new domains, locales or user inputs in a variety of different applications. The second oral session showcases several deployed systems. In addition to discussing the choices made for system architecture and standard evaluation metrics, these papers also report the impact on end users, the ultimate test of a system's usefulness. Work presented in the poster session paints a rich picture of the many real-world applications of NLP and speech technologies and the challenges associated with these applications.

NAACL-HLT 2019 Industry Track also features the “Careers in NLP” panel discussion. The rebranded edition of this panel discussion recognizes the diversity of NLP careers today. Traditional career paths have typically led NLP researchers into academia, industrial labs, and government agencies. Today, we also see an increase in roles at startup companies and an emerging NLP practitioner role in industry that intersects with software, data, and product. As last year, the panel will be moderated by Philip Resnik, professor at the University of Maryland, and we expect the conversation to include trends in NLP careers, emerging skills, prominent challenges and opportunities for cross-functional collaboration as NLP professionals in today’s organizations, and more.

The NAACL-HLT 2019 industry track program is the culmination of the small steps we have taken towards elevating and integrating this track further into the conference. We hope the program we have put together will strengthen the community’s resolve to continue to organize and attend a similar track at future conferences.

On a personal note, we recognize the privilege of chairing the NAACL-HLT 2019 Industry Track. We thank the conference general chair, Jill Burstein, for inviting us to the organizing committee. Thanks also to Program Chairs Christy, Ted and Tamar as well as all members of the organizing committee. We were generously helped by every member of this committee over the past year and organizing this track was possible only with their advice and efforts. We once again recognize and thank every member of the industry track program committee for volunteering their time. Finally, thanks to the authors and attendees of the industry track for embracing this initiative and offering a reason to continue the industry track at NAACL-HLT conferences.

Anastassia, Michelle, Rohit

Industry Track Co-chairs:

Rohit Kumar
Anastassia Loukina, Educational Testing Service
Michelle Morales, IBM

Program Committee:

Nitin Agarwal, Microsoft
Sachin Agarwal, Apple
Hua Ai, Delta Airlines
Alan Akbik, Zalando Research
Miguel Ballesteros, IBM Research AI
Nikoletta Basiou, SRI International
Frederic Bechet, Aix-Marseille Université
Trung Bui, Adobe Research
Donna Byron, IBM Cognitive Applications
Vitor Carvalho, Intuit AI
Francisco Casacuberta, Universitat Politècnica de València
Praveen Chandar, Spotify
Sourish Chaudhuri, Google Inc
Ciprian Chelba, Google
Wei Chen, Google
John Chen, Interactions LLC
Laura Chiticariu, IBM Watson
Justin Chiu, Avrio AI
Brooke Cowan, Expedia Group
Deborah Dahl, Conversational Technologies
Lingjia Deng, Bloomberg L.P.
Giuseppe Di Fabbrizio, VUI Inc.
Matthew Dunn, LivePerson
Keelan Evanini, Educational Testing Service
Oliver Ferschke, M*Modal
Michael Flor, Educational Testing Service
Rashmi Gangadharaiah, AWS AI Amazon
Anna Lisa, Gentile IBM Research Almaden
Rahul Goel, Amazon Alexa
Anuj Goyal, Amazon Alexa
Dilek Hakkani-Tur, Amazon Alexa AI
Sanjika Hewavitharana, eBay
Derrick Higgins, American Family Insurance
Lynette Hirschman, MITRE
Yufang Hou, IBM Research
Javid Huseynov, IBM
Rahul Jha, Microsoft Corporation
Mahesh Joshi, LinkedIn
Adi Kalyanpur, Elemental Cognition
Kartikay Khandelwal, Facebook
Saurabh Khanwalkar, Bose Corporation

Doo Soon Kim, Adobe Research
Sun Kim, National Center for Biotechnology Information (NCBI)
Jared Kramer, Amazon
Sanjeev Kumar, Quark.ai
Gakuto Kurata, IBM Research
Constantine Lignos, University of Southern California Information Sciences Institute
Bing Liu, University of Illinois at Chicago
Alexander Loeser, Beuth-University of Applied Sciences Berlin
Xiaoqiang Luo, LinkedIn
Nitin Madnani, Educational Testing Service
Arindam Mandal, Amazon.com
Yuji Matsumoto, NAIST / Riken AIP
Florian Metze, Carnegie Mellon University
Lisa Michaud, Aspect Software
Isabelle Moulinier, Capital One
Matthew Mulholland, Educational Testing Service
Udhyakumar Nallasamy, Apple
Elnaz Nouri, Microsoft Research
Mari Olsen, Lionbridge - Machine Intelligence
Stefano Pacifico, Epistemic AI
Aasish Pappu, Spotify Research
Youngja Park, IBM T. J. Watson Research Center
Ioannis Partalas, Expedia Group
Siddharth Patwardhan, Apple
Alexandros Potamianos, Behavioral Signals
John Prager, IBM Research
Rashmi Prasad, Interactions Corporation
Long Qin, Singsound Inc
Elio Querze, Bose Corp
Owen Rambow, Elemental Cognition LLC
Meghana Ravikumar, SigOpt
Sravana Reddy, Spotify
Ehud Reiter, University of Aberdeen
Steve Renals, University of Edinburgh
Giuseppe Riccardi, University of Trento
Brian Riordan, Educational Testing Service
Salim Roukos, IBM
Nicholas Ruiz, Interactions LLC
Alicia Sagae, Consulting Scientist
Avneesh Saluja, Netflix
Stefan Scherer, Embodied Inc.
Frank Schilder, Thomson Reuters
Frank Seide, Microsoft Research
Ethan Selfridge, Interactions LLC
Rushin Shah, Facebook
Michal Shmueli-Scheuer, IBM Research
Sunayana Sitaram, Microsoft Research India
Kazoo Sone, Google
Biplav Srivastava, IBM
Svetlana Stoyanchev, Interactions Corporation
David Suendermann-Oeft, Modality

Isabel Trancoso, INESC-ID / IST
Keith Trnka, 98point6 Inc.
Ling Tsou, SDL
Gokhan Tur, Uber
Ngoc Phuoc An Vo, IBM Research
Xinhao Wang, Educational Testing Service
Yi-Chia Wang, Uber AI
Jason D Williams, Apple

Careers in NLP Panel:

Judith L. Klavans, Independent
Yun Yao Li, IBM Research
Owen Rambow, Elemental Cognition
Philip Resnik (Moderator), University of Maryland
Joel Tetreault, Grammarly

Table of Contents

<i>Enabling Real-time Neural IME with Incremental Vocabulary Selection</i> Jiali Yao, Raphael Shu, Xinjian Li, Katsutoshi Ohtsuki and Hideki Nakayama	1
<i>Locale-agnostic Universal Domain Classification Model in Spoken Language Understanding</i> Jihwan Lee, Ruhi Sarikaya and Young-Bum Kim	9
<i>Practical Semantic Parsing for Spoken Language Understanding</i> Marco Damonte, Rahul Goel and Tagyoung Chung	16
<i>Fast Prototyping a Dialogue Comprehension System for Nurse-Patient Conversations on Symptom Monitoring</i> Zhengyuan Liu, Hazel Lim, Nur Farah Ain Suhaimi, Shao Chuen Tong, Sharon Ong, Angela Ng, Sheldon Lee, Michael R. Macdonald, Savitha Ramasamy, Pavitra Krishnaswamy, Wai Leng Chow and Nancy F. Chen	24
<i>Graph Convolution for Multimodal Information Extraction from Visually Rich Documents</i> Xiaojing Liu, Feiyu Gao, Qiong Zhang and Huasha Zhao	32
<i>Diversifying Reply Suggestions Using a Matching-Conditional Variational Autoencoder</i> Budhaditya Deb, Peter Bailey and Milad Shokouhi	40
<i>Goal-Oriented End-to-End Conversational Models with Profile Features in a Real-World Setting</i> Yichao Lu, Manisha Srivastava, Jared Kramer, Heba Elfardy, Andrea Kahn, Song Wang and Vikas Bhardwaj	48
<i>Detecting Customer Complaint Escalation with Recurrent Neural Networks and Manually-Engineered Features</i> Wei Yang, Luchen Tan, Chunwei Lu, Anqi Cui, Han Li, Xi Chen, Kun Xiong, Muzi Wang, Ming Li, Jian Pei and Jimmy Lin	56
<i>Multi-Modal Generative Adversarial Network for Short Product Title Generation in Mobile E-Commerce</i> Jianguo Zhang, Pengcheng Zou, Zhao Li, Yao Wan, Xiuming Pan, Yu Gong and Philip S. Yu ...	64
<i>A Case Study on Neural Headline Generation for Editing Support</i> Kazuma Murao, Ken Kobayashi, Hayato Kobayashi, Taichi Yatsuka, Takeshi Masuyama, Tatsuru Higurashi and Yoshimune Tabuchi	73
<i>Neural Lexicons for Slot Tagging in Spoken Language Understanding</i> Kyle Williams	83
<i>Active Learning for New Domains in Natural Language Understanding</i> Stanislav Peshterliev, John Kearney, Abhyuday Jagannatha, Imre Kiss and Spyros Matsoukas ..	90
<i>Scaling Multi-Domain Dialogue State Tracking via Query Reformulation</i> Pushpendre Rastogi, Arpit Gupta, Tongfei Chen and Mathias Lambert	97
<i>Are the Tools up to the Task? an Evaluation of Commercial Dialog Tools in Developing Conversational Enterprise-grade Dialog Systems</i> Marie Meteer, Meghan Hickey, Carmi Rothberg, David Nahamoo and Ellen Eide Kislal	106

<i>Development and Deployment of a Large-Scale Dialog-based Intelligent Tutoring System</i> Shazia Afzal, Tejas Dhamecha, Nirmal Mukhi, Renuka Sindhgatta, Smit Marvaniya, Matthew Ventura and Jessica Yarbrow	114
<i>Learning When Not to Answer: a Ternary Reward Structure for Reinforcement Learning Based Question Answering</i> Frédéric Godin, Anjishnu Kumar and Arpit Mittal	122
<i>Extraction of Message Sequence Charts from Software Use-Case Descriptions</i> Girish Palshikar, Nitin Ramrakhiyani, Sangameshwar Patil, Sachin Pawar, Swapnil Hingmire, Vasudeva Varma and Pushpak Bhattacharyya	130
<i>Improving Knowledge Base Construction from Robust Infobox Extraction</i> Boya Peng, Yejin Huh, Xiao Ling and Michele Banko	138
<i>A k-Nearest Neighbor Approach towards Multi-level Sequence Labeling</i> Yue Chen and John Chen	149
<i>Train One Get One Free: Partially Supervised Neural Network for Bug Report Duplicate Detection and Clustering</i> Lahari Poddar, Leonardo Neves, William Brendel, Luis Marujo, Sergey Tulyakov and Pradeep Karaturi	157
<i>Robust Semantic Parsing with Adversarial Learning for Domain Generalization</i> Gabriel Marzinotto, Geraldine Damnati, Frederic Bechet and Benoit Favre	166
<i>TOI-CNN: a Solution of Information Extraction on Chinese Insurance Policy</i> Lin Sun, Kai Zhang, Fule Ji and Zhenhua Yang	174
<i>Cross-lingual Transfer Learning for Japanese Named Entity Recognition</i> Andrew Johnson, Penny Karanasou, Judith Gaspers and Dietrich Klakow	182
<i>Neural Text Normalization with Subword Units</i> Courtney Mansfield, Ming Sun, Yuzong Liu, Ankur Gandhe and Bjorn Hoffmeister	190
<i>Audio De-identification - a New Entity Recognition Task</i> Ido Cohn, Itay Laish, Genady Beryozkin, Gang Li, Izhak Shafran, Idan Szpektor, Tzvika Hartman, Avinatan Hassidim and Yossi Matias	197
<i>In Other News: a Bi-style Text-to-speech Model for Synthesizing Newscaster Voice with Limited Data</i> Nishant Prateek, Mateusz Łajszczak, Roberto Barra-Chicote, Thomas Drugman, Jaime Lorenzo-Trueba, Thomas Merritt, Srikanth Ronanki and Trevor Wood	205
<i>Generate, Filter, and Rank: Grammaticality Classification for Production-Ready NLG Systems</i> Ashwini Challa, Kartikeya Upasani, Anusha Balakrishnan and Rajen Subba	214
<i>Content-based Dwell Time Engagement Prediction Model for News Articles</i> Heidar Davoudi, Aijun An and Gordon Edall	226

Conference Program

Monday, June 3, 2019

13:00–14:30 *Careers in NLP Panel*

Tuesday June 4, 2019

09:00–10:30 *Oral Sessions (long papers) and Posters (long and short papers)*

Session 4E (Industry): Real world challenges

09:00–09:18 *Enabling Real-time Neural IME with Incremental Vocabulary Selection*
Jiali Yao, Raphael Shu, Xinjian Li, Katsutoshi Ohtsuki and Hideki Nakayama

09:18–09:36 *Locale-agnostic Universal Domain Classification Model in Spoken Language Understanding*
Jihwan Lee, Ruhi Sarikaya and Young-Bum Kim

09:36–09:54 *Practical Semantic Parsing for Spoken Language Understanding*
Marco Damonte, Rahul Goel and Tagyoung Chung

09:54–10:12 *Fast Prototyping a Dialogue Comprehension System for Nurse-Patient Conversations on Symptom Monitoring*
Zhengyuan Liu, Hazel Lim, Nur Farah Ain Suhaimi, Shao Chuen Tong, Sharon Ong, Angela Ng, Sheldon Lee, Michael R. Macdonald, Savitha Ramasamy, Pavitra Krishnaswamy, Wai Leng Chow and Nancy F. Chen

10:12–10:30 *Graph Convolution for Multimodal Information Extraction from Visually Rich Documents*
Xiaojing Liu, Feiyu Gao, Qiong Zhang and Huasha Zhao

15:30–17:00 *Oral sessions (long papers) and Posters (long and short papers)*

Tuesday June 4, 2019 (continued)

Session 6E (Industry): Deployed systems

- 15:30–15:48 *Diversifying Reply Suggestions Using a Matching-Conditional Variational Autoencoder*
Budhaditya Deb, Peter Bailey and Milad Shokouhi
- 15:48–16:06 *Goal-Oriented End-to-End Conversational Models with Profile Features in a Real-World Setting*
Yichao Lu, Manisha Srivastava, Jared Kramer, Heba Elfardy, Andrea Kahn, Song Wang and Vikas Bhardwaj
- 16:06–16:24 *Detecting Customer Complaint Escalation with Recurrent Neural Networks and Manually-Engineered Features*
Wei Yang, Luchen Tan, Chunwei Lu, Anqi Cui, Han Li, Xi Chen, Kun Xiong, Muzi Wang, Ming Li, Jian Pei and Jimmy Lin
- 16:24–16:42 *Multi-Modal Generative Adversarial Network for Short Product Title Generation in Mobile E-Commerce*
Jianguo Zhang, Pengcheng Zou, Zhao Li, Yao Wan, Xiuming Pan, Yu Gong and Philip S. Yu
- 16:42–17:00 *A Case Study on Neural Headline Generation for Editing Support*
Kazuma Murao, Ken Kobayashi, Hayato Kobayashi, Taichi Yatsuka, Takeshi Masuyama, Tatsuru Higurashi and Yoshimune Tabuchi

Wednesday June 5, 2019

15:30–16:30 Oral Sessions (short papers) and Posters (Industry)

Session 9F: Industry posters

- 15:30–16:30 *Neural Lexicons for Slot Tagging in Spoken Language Understanding*
Kyle Williams
- 15:30–16:30 *Active Learning for New Domains in Natural Language Understanding*
Stanislav Peshterliev, John Kearney, Abhyuday Jagannatha, Imre Kiss and Spyros Matsoukas
- 15:30–16:30 *Scaling Multi-Domain Dialogue State Tracking via Query Reformulation*
Pushpendre Rastogi, Arpit Gupta, Tongfei Chen and Mathias Lambert
- 15:30–16:30 *Are the Tools up to the Task? an Evaluation of Commercial Dialog Tools in Developing Conversational Enterprise-grade Dialog Systems*
Marie Meteer, Meghan Hickey, Carmi Rothberg, David Nahamoo and Ellen Eide Kislal

Wednesday June 5, 2019 (continued)

- 15:30–16:30 *Development and Deployment of a Large-Scale Dialog-based Intelligent Tutoring System*
Shazia Afzal, Tejas Dhamecha, Nirmal Mukhi, Renuka Sindhgatta, Smit Marvaniya, Matthew Ventura and Jessica Yarbrow
- 15:30–16:30 *Learning When Not to Answer: a Ternary Reward Structure for Reinforcement Learning Based Question Answering*
Frédéric Godin, Anjishnu Kumar and Arpit Mittal
- 15:30–16:30 *Extraction of Message Sequence Charts from Software Use-Case Descriptions*
Girish Palshikar, Nitin Ramrakhiyani, Sangameshwar Patil, Sachin Pawar, Swapnil Hingmire, Vasudeva Varma and Pushpak Bhattacharyya
- 15:30–16:30 *Improving Knowledge Base Construction from Robust Infobox Extraction*
Boya Peng, Yejin Huh, Xiao Ling and Michele Banko
- 15:30–16:30 *A k-Nearest Neighbor Approach towards Multi-level Sequence Labeling*
Yue Chen and John Chen
- 15:30–16:30 *Train One Get One Free: Partially Supervised Neural Network for Bug Report Duplicate Detection and Clustering*
Lahari Poddar, Leonardo Neves, William Brendel, Luis Marujo, Sergey Tulyakov and Pradeep Karuturi
- 15:30–16:30 *Robust Semantic Parsing with Adversarial Learning for Domain Generalization*
Gabriel Marzinotto, Geraldine Damnati, Frederic Bechet and Benoit Favre
- 15:30–16:30 *TOI-CNN: a Solution of Information Extraction on Chinese Insurance Policy*
Lin Sun, Kai Zhang, Fule Ji and Zhenhua Yang
- 15:30–16:30 *Cross-lingual Transfer Learning for Japanese Named Entity Recognition*
Andrew Johnson, Penny Karanasou, Judith Gaspers and Dietrich Klakow
- 15:30–16:30 *Neural Text Normalization with Subword Units*
Courtney Mansfield, Ming Sun, Yuzong Liu, Ankur Gandhe and Bjorn Hoffmeister
- 15:30–16:30 *Audio De-identification - a New Entity Recognition Task*
Ido Cohn, Itay Laish, Genady Beryozkin, Gang Li, Izhak Shafran, Idan Szpektor, Tzvika Hartman, Avinatan Hassidim and Yossi Matias
- 15:30–16:30 *In Other News: a Bi-style Text-to-speech Model for Synthesizing Newscaster Voice with Limited Data*
Nishant Prateek, Mateusz Łajszczak, Roberto Barra-Chicote, Thomas Drugman, Jaime Lorenzo-Trueba, Thomas Merritt, Srikanth Ronanki and Trevor Wood

Wednesday June 5, 2019 (continued)

15:30–16:30 *Generate, Filter, and Rank: Grammaticality Classification for Production-Ready NLG Systems*

Ashwini Challa, Kartikeya Upasani, Anusha Balakrishnan and Rajen Subba

15:30–16:30 *Content-based Dwell Time Engagement Prediction Model for News Articles*

Heidar Davoudi, Aijun An and Gordon Edall