

W-NUT 2019

**The Fifth Workshop on  
Noisy User-generated Text  
(W-NUT 2019)**

**Proceedings of the Workshop**

Nov 4, 2019  
Hong Kong, China

©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](mailto:acl@aclweb.org)

ISBN 978-1-950737-84-0

## **Introduction**

The W-NUT 2019 workshop focuses on a core set of natural language processing tasks on top of noisy user-generated text, such as that found on social media, web forums and online reviews. Recent years have seen a significant increase of interest in these areas. The internet has democratized content creation leading to an explosion of informal user-generated text, publicly available in electronic format, motivating the need for NLP on noisy text to enable new data analytics applications.

We received 89 long and short paper submissions this year. There are two invited speakers, Isabelle Augenstein (University of Copenhagen) and Jing Jiang (Singapore Management University) with each of their talks covering a different aspect of NLP for user-generated text. We have the best paper award(s) sponsored by Google this year, for which we are thankful. We would like to thank the Program Committee members who reviewed the papers this year. We would also like to thank the workshop participants.

Wei Xu, Alan Ritter, Tim Baldwin and Afshin Rahimi  
Co-Organizers



**Organizers:**

Wei Xu, Ohio State University  
Alan Ritter, Ohio State University  
Tim Baldwin, University of Melbourne  
Afshin Rahimi, University of Melbourne

**Program Committee:**

Mostafa Abdou (University of Copenhagen)  
Muhammad Abdul-Mageed (University of British Columbia)  
Željko Agić (Corti)  
Gustavo Aguilar (University of Houston)  
Hadi Amiri (Harvard University)  
Rahul Aralikatte (University of Copenhagen)  
Eiji Aramaki (NAIST)  
Roy Bar-Haim (IBM)  
Francesco Barbieri (UPF Barcelona)  
Cosmin Bejan (Vanderbilt University)  
Eric Bell (PNNL)  
Adrian Benton (JHU)  
Eduardo Blanco (University of North Texas)  
Su Lin Blodgett (UMass Amherst)  
Matko Bošnjak (University College London)  
Julian Brooke (University of British Columbia)  
Annabelle Carrell (JHU)  
Xilun Chen (Cornell University)  
Anne Cocos (University of Pennsylvania)  
Arman Cohan (AI2)  
Nigel Collier (University of Cambridge)  
Paul Cook (University of New Brunswick)  
Marina Danilevsky (IBM Research)  
Leon Derczynski (IT University of Copenhagen)  
Seza Dođruöz (Tilburg University)  
Jay DeYoung (Northeastern University)  
Eduard Dragut (Temple University)  
Xinya Du (Cornell University)  
Heba Elfardy (Amazon)  
Micha Elsner (Ohio State University)  
Sindhu Kiranmai Ernala (Georgia Tech)  
Manaal Faruqui (Google Research)  
Lisheng Fu (New York University)  
Yoshinari Fujinuma (University of Colorado, Boulder)  
Dan Garrette (Google Research)  
Kevin Gimpel (TTIC)  
Dan Goldwasser (Purdue University)  
Amit Goyal (Criteo)  
Nizar Habash (NYU Abu Dhabi)  
Masato Hagiwara (Duolingo)

Bo Han (Kaplan)  
Abe Handler (University of Massachusetts Amherst)  
Shudong Hao (University of Colorado, Boulder)  
Devamanyu Hazarika (National University of Singapore)  
Jack Hessel (Cornell University)  
Dirk Hovy (Bocconi University)  
Xiaolei Huang (University of Colorado, Boulder)  
Sarthak Jain (Northeastern University)  
Kenny Joseph (University at Buffalo)  
David Jurgens (University of Michigan)  
Nobuhiro Kaji (Yahoo! Research)  
Pallika Kanani (Oracle)  
Dongyeop Kang (Carnegie Mellon University)  
Emre Kiciman (Microsoft Research)  
Svetlana Kiritchenko (National Research Council Canada)  
Roman Klinger (University of Stuttgart)  
Ekaterina Kochmar (University of Cambridge)  
Vivek Kulkarni (University of California Santa Barbara)  
Jonathan Kummerfeld (University of Michigan)  
Ophélie Lacroix (Siteimprove)  
Wuwei Lan (Ohio State University)  
Chen Li (Tencent)  
Jing Li (Tencent AI)  
Jessy Junyi Li (University of Texas Austin)  
Yitong Li (University of Melbourne)  
Nut Limsopatham (University of Glasgow)  
Patrick Littell (National Research Council Canada)  
Zhiyuan Liu (Tsinghua University)  
Fei Liu (University of Melbourne)  
Nikola Ljubešić (University of Zagreb)  
Wei-Yun Ma (Academia Sinica)  
Mounica Maddela (Ohio State University)  
Suraj Maharjan (University of Houston)  
Aaron Masino (The Children's Hospital of Philadelphia)  
Paul Michel (CMU)  
Shachar Mirkin (Xerox Research)  
Saif M. Mohammad (National Research Council Canada)  
Ahmed Mourad (RMIT University)  
Günter Neumann (DFKI)  
Vincent Ng (University of Texas at Dallas)  
Eric Nichols (Honda Research Institute)  
Xing Niu (University of Maryland, College Park)  
Benjamin Nye (Northeastern University)  
Alice Oh (KAIST)  
Naoki Otani (CMU)  
Patrick Pantel (Microsoft Research)  
Umashanthi Pavalanathan (Georgia Tech)  
Yuval Pinter (Georgia Tech)  
Barbara Plank (IT University of Copenhagen)  
Christopher Potts (Stanford University)  
Daniel Preoțiuc-Pietro (Bloomberg)

Chris Quirk (Microsoft Research)  
Ella Rabinovich (University of Toronto)  
Dianna Radpour (University of Colorado Boulder)  
Preethi Raghavan (IBM Research)  
Revanth Rameshkumar (Microsoft)  
Sudha Rao (Microsoft Research)  
Marek Rei (University of Cambridge)  
Roi Reichart (Technion)  
Adithya Renduchintala (JHU)  
Carolyn Penstein Rose (CMU)  
Alla Rozovskaya (City University of New York)  
Koustuv Saha (Georgia Tech)  
Keisuke Sakaguchi (Allen Institute for Artificial Intelligence)  
Maarten Sap (University of Washington)  
Natalie Schluter (IT University of Copenhagen)  
Andrew Schwartz (Stony Brook University)  
Djamé Seddah (University Paris-Sorbonne)  
Amirreza Shirani (University of Houston)  
Dan Simonson (BlackBoiler)  
Evangelia Spiliopoulou (Carnegie Mellon University)  
Jan Šnajder (University of Zagreb)  
Gabriel Stanovsky (Allen Institute for Artificial Intelligence)  
Ian Stewart (Georgia Tech)  
Jeniya Tabassum (Ohio State University)  
Joel Tetreault (Grammarly)  
Sara Tonelli (FBK)  
Rob van der Goot (University of Groningen)  
Rob Voigt (Stanford University)  
Byron Wallace (Northeastern University)  
Xiaojun Wan (Peking University)  
Zeerak Waseem (University of Sheffield)  
Zhongyu Wei (Fudan University)  
Diyi Yang (Georgia Tech)  
Yi Yang (ASAPP)  
Guido Zarrella (MITRE)  
Justine Zhang (Cornell University)  
Jason Shuo Zhang (University of Colorado, Boulder)  
Shi Zong (Ohio State University)

**Invited Speakers:**

Isabelle Augenstein (University of Copenhagen)  
Jing Jiang (Singapore Management University)





## Table of Contents

<i>Weakly Supervised Attention Networks for Fine-Grained Opinion Mining and Public Health</i> Giannis Karamanolakis, Daniel Hsu and Luis Gravano .....	1
<i>Formality Style Transfer for Noisy, User-generated Conversations: Extracting Labeled, Parallel Data from Unlabeled Corpora</i> Isak Czeresnia Etinger and Alan W Black .....	11
<i>Multilingual Whispers: Generating Paraphrases with Translation</i> Christian Federmann, Oussama Elachqar and Chris Quirk .....	17
<i>Personalizing Grammatical Error Correction: Adaptation to Proficiency Level and L1</i> Maria Nadejde and Joel Tetreault .....	27
<i>Exploiting BERT for End-to-End Aspect-based Sentiment Analysis</i> Xin Li, Lidong Bing, Wenxuan Zhang and Wai Lam .....	34
<i>Training on Synthetic Noise Improves Robustness to Natural Noise in Machine Translation</i> vladimir karpukhin, Omer Levy, Jacob Eisenstein and Marjan Ghazvininejad .....	42
<i>Character-Based Models for Adversarial Phone Extraction: Preventing Human Sex Trafficking</i> Nathanael Chambers, Timothy Forman, Catherine Griswold, Kevin Lu, Yogoish Khastgir and Stephen Steckler .....	48
<i>Tkol, Httt, and r/radiohead: High Affinity Terms in Reddit Communities</i> Abhinav Bhandari and Caitrin Armstrong .....	57
<i>Large Scale Question Paraphrase Retrieval with Smoothed Deep Metric Learning</i> Daniele Bonadiman, Anjishnu Kumar and Arpit Mittal .....	68
<i>Hey Siri. Ok Google. Alexa: A topic modeling of user reviews for smart speakers</i> Hanh Nguyen and Dirk Hovy .....	76
<i>Predicting Algorithm Classes for Programming Word Problems</i> vinayak athavale, aayush naik, rajas vanjape and Manish Shrivastava .....	84
<i>Automatic identification of writers' intentions: Comparing different methods for predicting relationship goals in online dating profile texts</i> Chris van der Lee, Tess van der Zanden, Emiel Kraemer, Maria Mos and Alexander Schouten ..	94
<i>Contextualized Word Representations from Distant Supervision with and for NER</i> Abbas Ghaddar and Phillippe Langlais .....	101
<i>Extract, Transform and Filling: A Pipeline Model for Question Paraphrasing based on Template</i> Yunfan Gu, yang yuqiao and Zhongyu Wei .....	109
<i>An In-depth Analysis of the Effect of Lexical Normalization on the Dependency Parsing of Social Media</i> Rob van der Goot .....	115
<i>Who wrote this book? A challenge for e-commerce</i> Béranger Dumont, Simona Maggio, Ghiles Sidi Said and Quoc-Tien Au .....	121
<i>Mining Tweets that refer to TV programs with Deep Neural Networks</i> Takeshi Kobayakawa, Taro Miyazaki, Hiroki Okamoto and Simon Clippingdale .....	126

<i>Normalising Non-standardised Orthography in Algerian Code-switched User-generated Data</i> Wafia Adouane, Jean-Philippe Bernardy and Simon Dobnik .....	131
<i>Dialect Text Normalization to Normative Standard Finnish</i> Niko Partanen, Mika Hämäläinen and Khalid Alnajjar .....	141
<i>A Cross-Topic Method for Supervised Relevance Classification</i> Jiawei Yong .....	147
<i>Exploring Multilingual Syntactic Sentence Representations</i> Chen Liu, Anderson De Andrade and Muhammad Osama .....	153
<i>FASpell: A Fast, Adaptable, Simple, Powerful Chinese Spell Checker Based On DAE-Decoder Paradigm</i> Yuzhong Hong, Xianguo Yu, Neng He, Nan Liu and Junhui Liu .....	160
<i>Latent semantic network induction in the context of linked example senses</i> Hunter Heidenreich and Jake Williams .....	170
<i>SmokEng: Towards Fine-grained Classification of Tobacco-related Social Media Text</i> Kartikey Pant, Venkata Himakar Yanamandra, Alok Debnath and Radhika Mamidi .....	181
<i>Modelling Uncertainty in Collaborative Document Quality Assessment</i> Aili Shen, Daniel Beck, Bahar Salehi, Jianzhong Qi and Timothy Baldwin .....	191
<i>Conceptualisation and Annotation of Drug Nonadherence Information for Knowledge Extraction from Patient-Generated Texts</i> Anja Belz, Richard Hoile, Elizabeth Ford and Azam Mullick .....	202
<i>Dataset Analysis and Augmentation for Emoji-Sensitive Irony Detection</i> Shirley Anugrah Hayati, Aditi Chaudhary, Naoki Otani and Alan W Black .....	212
<i>Geolocation with Attention-Based Multitask Learning Models</i> Tommaso Fornaciari and Dirk Hovy .....	217
<i>Dense Node Representation for Geolocation</i> Tommaso Fornaciari and Dirk Hovy .....	224
<i>Identifying Linguistic Areas for Geolocation</i> Tommaso Fornaciari and Dirk Hovy .....	231
<i>Robustness to Capitalization Errors in Named Entity Recognition</i> Sravan Bodapati, Hyokun Yun and Yaser Al-Onaizan .....	237
<i>Extending Event Detection to New Types with Learning from Keywords</i> Viet Dac Lai and Thien Nguyen .....	243
<i>Distant Supervised Relation Extraction with Separate Head-Tail CNN</i> Rui Xing and Jie Luo .....	249
<i>Discovering the Functions of Language in Online Forums</i> Youmna Ismaeil, Oana Balalau and Paramita Mirza .....	259
<i>Incremental processing of noisy user utterances in the spoken language understanding task</i> Stefan Constantin, Jan Niehues and Alex Waibel .....	265

<i>Benefits of Data Augmentation for NMT-based Text Normalization of User-Generated Content</i> Claudia Matos Veliz, Orphee De Clercq and Veronique Hoste .....	275
<i>Contextual Text Denoising with Masked Language Model</i> Yifu Sun and Haoming Jiang .....	286
<i>Towards Automated Semantic Role Labelling of Hindi-English Code-Mixed Tweets</i> Riya Pal and Dipti Sharma .....	291
<i>Enhancing BERT for Lexical Normalization</i> Benjamin Muller, Benoit Sagot and Djamé Seddah .....	297
<i>No, you're not alone: A better way to find people with similar experiences on Reddit</i> Zhilin Wang, Elena Rastorgueva, Weizhe Lin and Xiaodong Wu .....	307
<i>Improving Multi-label Emotion Classification by Integrating both General and Domain-specific Knowledge</i> Wenhao Ying, Rong Xiang and Qin Lu .....	316
<i>Adapting Deep Learning Methods for Mental Health Prediction on Social Media</i> Ivan Sekulic and Michael Strube .....	322
<i>Improving Neural Machine Translation Robustness via Data Augmentation: Beyond Back-Translation</i> Zhenhao Li and Lucia Specia .....	328
<i>An Ensemble of Humour, Sarcasm, and Hate Speech for Sentiment Classification in Online Reviews</i> Rohan Badlani, Nishit Asnani and Manan Rai .....	337
<i>Grammatical Error Correction in Low-Resource Scenarios</i> Jakub Náplava and Milan Straka .....	346
<i>Minimally-Augmented Grammatical Error Correction</i> Roman Grundkiewicz and Marcin Junczys-Dowmunt .....	357
<i>A Social Opinion Gold Standard for the Malta Government Budget 2018</i> Keith Cortis and Brian Davis .....	364
<i>The Fallacy of Echo Chambers: Analyzing the Political Slants of User-Generated News Comments in Korean Media</i> Jiyoung Han, Youngin Lee, Junbum Lee and Meeyoung Cha .....	370
<i>Y'all should read this! Identifying Plurality in Second-Person Personal Pronouns in English Texts</i> Gabriel Stanovsky and Ronen Tamari .....	375
<i>An Edit-centric Approach for Wikipedia Article Quality Assessment</i> Edison Marrese-Taylor, Pablo Loyola and Yutaka Matsuo .....	381
<i>Additive Compositionality of Word Vectors</i> Yeon Seonwoo, Sungjoon Park, Dongkwan Kim and Alice Oh .....	387
<i>Contextualized context2vec</i> Kazuki Ashihara, Tomoyuki Kajiwara, Yuki Arase and Satoru Uchida .....	397
<i>Phonetic Normalization for Machine Translation of User Generated Content</i> José Carlos Rosales Núñez, Djamé Seddah and Guillaume Wisniewski .....	407

<i>Normalization of Indonesian-English Code-Mixed Twitter Data</i>	
Anab Maulana Barik, Rahmad Mahendra and Mirna Adriani .....	417
<i>Unsupervised Neologism Normalization Using Embedding Space Mapping</i>	
Nasser Zalmout, Kapil Thadani and Aasish Pappu .....	425
<i>Lexical Features Are More Vulnerable, Syntactic Features Have More Predictive Power</i>	
Jekaterina Novikova, Aparna Balagopalan, Ksenia Shkaruta and Frank Rudzicz .....	431
<i>Towards Actual (Not Operational) Textual Style Transfer Auto-Evaluation</i>	
Richard Yuanzhe Pang .....	444
<i>CodeSwitch-Reddit: Exploration of Written Multilingual Discourse in Online Discussion Forums</i>	
Ella Rabinovich, Masih Sultani and Suzanne Stevenson .....	446

# Conference Program

Monday, November, 4, 2019

**9:00–9:05**      **Opening**

**9:05–9:50**      **Invited Talk: Isabelle Augenstein**

**9:50–10:35**      **Oral Session I**

9:50–10:05      *Weakly Supervised Attention Networks for Fine-Grained Opinion Mining and Public Health*

Giannis Karamanolakis, Daniel Hsu and Luis Gravano

10:05–10:20      *Formality Style Transfer for Noisy, User-generated Conversations: Extracting Labeled, Parallel Data from Unlabeled Corpora*

Isak Czeresnia Etinger and Alan W Black

10:20–10:35      *Multilingual Whispers: Generating Paraphrases with Translation*

Christian Federmann, Oussama Elachqar and Chris Quirk

**10:35–11:00**      **Coffee Break**

**11:00–12:15**      **Oral Session II**

11:00–11:15      *Personalizing Grammatical Error Correction: Adaptation to Proficiency Level and L1*

Maria Nadejde and Joel Tetreault

11:15–11:30      *Exploiting BERT for End-to-End Aspect-based Sentiment Analysis*

Xin Li, Lidong Bing, Wenxuan Zhang and Wai Lam

11:30–11:45      *Training on Synthetic Noise Improves Robustness to Natural Noise in Machine Translation*

vladimir karpukhin, Omer Levy, Jacob Eisenstein and Marjan Ghazvininejad

11:45–12:00      *Character-Based Models for Adversarial Phone Extraction: Preventing Human Sex Trafficking*

Nathanael Chambers, Timothy Forman, Catherine Griswold, Kevin Lu, Yogaish Khastgir and Stephen Steckler

**Monday, November, 4, 2019 (continued)**

12:00–12:15 *Tkol, Httt, and r/radiohead: High Affinity Terms in Reddit Communities*  
Abhinav Bhandari and Caitrin Armstrong

**12:30–2:00 Lunch**

**2:00–3:00 Lightning Talks**

*Large Scale Question Paraphrase Retrieval with Smoothed Deep Metric Learning*  
Daniele Bonadiman, Anjishnu Kumar and Arpit Mittal

*Hey Siri. Ok Google. Alexa: A topic modeling of user reviews for smart speakers*  
Hanh Nguyen and Dirk Hovy

*Predicting Algorithm Classes for Programming Word Problems*  
vinayak athavale, aayush naik, rajas vanjape and Manish Shrivastava

*Automatic identification of writers' intentions: Comparing different methods for predicting relationship goals in online dating profile texts*  
Chris van der Lee, Tess van der Zanden, Emiel Krahmer, Maria Mos and Alexander Schouten

*Contextualized Word Representations from Distant Supervision with and for NER*  
Abbas Ghaddar and Phillippe Langlais

*Extract, Transform and Filling: A Pipeline Model for Question Paraphrasing based on Template*  
Yunfan Gu, yang yuqiao and Zhongyu Wei

*An In-depth Analysis of the Effect of Lexical Normalization on the Dependency Parsing of Social Media*  
Rob van der Goot

*Who wrote this book? A challenge for e-commerce*  
Béranger Dumont, Simona Maggio, Ghiles Sidi Said and Quoc-Tien Au

*Mining Tweets that refer to TV programs with Deep Neural Networks*  
Takeshi Kobayakawa, Taro Miyazaki, Hiroki Okamoto and Simon Clippingdale

**Monday, November, 4, 2019 (continued)**

*Normalising Non-standardised Orthography in Algerian Code-switched User-generated Data*

Wafia Adouane, Jean-Philippe Bernardy and Simon Dobnik

*Dialect Text Normalization to Normative Standard Finnish*

Niko Partanen, Mika Hämäläinen and Khalid Alnajjar

*A Cross-Topic Method for Supervised Relevance Classification*

Jiawei Yong

*Exploring Multilingual Syntactic Sentence Representations*

Chen Liu, Anderson De Andrade and Muhammad Osama

*FASpell: A Fast, Adaptable, Simple, Powerful Chinese Spell Checker Based On DAE-Decoder Paradigm*

Yuzhong Hong, Xianguo Yu, Neng He, Nan Liu and Junhui Liu

*Latent semantic network induction in the context of linked example senses*

Hunter Heidenreich and Jake Williams

*SmokEng: Towards Fine-grained Classification of Tobacco-related Social Media Text*

Kartikey Pant, Venkata Himakar Yanamandra, Alok Debnath and Radhika Mamidi

*Modelling Uncertainty in Collaborative Document Quality Assessment*

Aili Shen, Daniel Beck, Bahar Salehi, Jianzhong Qi and Timothy Baldwin

*Conceptualisation and Annotation of Drug Nonadherence Information for Knowledge Extraction from Patient-Generated Texts*

Anja Belz, Richard Hoile, Elizabeth Ford and Azam Mullick

*Dataset Analysis and Augmentation for Emoji-Sensitive Irony Detection*

Shirley Anugrah Hayati, Aditi Chaudhary, Naoki Otani and Alan W Black

*Geolocation with Attention-Based Multitask Learning Models*

Tommaso Fornaciari and Dirk Hovy

*Dense Node Representation for Geolocation*

Tommaso Fornaciari and Dirk Hovy

**Monday, November, 4, 2019 (continued)**

*Identifying Linguistic Areas for Geolocation*

Tommaso Fornaciari and Dirk Hovy

*Robustness to Capitalization Errors in Named Entity Recognition*

Sravan Bodapati, Hyokun Yun and Yaser Al-Onaizan

*Extending Event Detection to New Types with Learning from Keywords*

Viet Dac Lai and Thien Nguyen

*Distant Supervised Relation Extraction with Separate Head-Tail CNN*

Rui Xing and Jie Luo

*Discovering the Functions of Language in Online Forums*

Youmna Ismaeil, Oana Balalau and Paramita Mirza

*Incremental processing of noisy user utterances in the spoken language understanding task*

Stefan Constantin, Jan Niehues and Alex Waibel

*Benefits of Data Augmentation for NMT-based Text Normalization of User-Generated Content*

Claudia Matos Veliz, Orphee De Clercq and Veronique Hoste

*Contextual Text Denoising with Masked Language Model*

Yifu Sun and Haoming Jiang

*Towards Automated Semantic Role Labelling of Hindi-English Code-Mixed Tweets*

Riya Pal and Dipti Sharma

*Enhancing BERT for Lexical Normalization*

Benjamin Muller, Benoit Sagot and Djamé Seddah

*No, you're not alone: A better way to find people with similar experiences on Reddit*

Zhilin Wang, Elena Rastorgueva, Weizhe Lin and Xiaodong Wu

*Improving Multi-label Emotion Classification by Integrating both General and Domain-specific Knowledge*

Wenhao Ying, Rong Xiang and Qin Lu



**Monday, November, 4, 2019 (continued)**

*Adapting Deep Learning Methods for Mental Health Prediction on Social Media*

Ivan Sekulic and Michael Strube

*Improving Neural Machine Translation Robustness via Data Augmentation: Beyond Back-Translation*

Zhenhao Li and Lucia Specia

*An Ensemble of Humour, Sarcasm, and Hate Speech for Sentiment Classification in Online Reviews*

Rohan Badlani, Nishit Asnani and Manan Rai

*Grammatical Error Correction in Low-Resource Scenarios*

Jakub Náplava and Milan Straka

*Minimally-Augmented Grammatical Error Correction*

Roman Grundkiewicz and Marcin Junczys-Dowmunt

*A Social Opinion Gold Standard for the Malta Government Budget 2018*

Keith Cortis and Brian Davis

*The Fallacy of Echo Chambers: Analyzing the Political Slants of User-Generated News Comments in Korean Media*

Jiyoung Han, Youngin Lee, Junbum Lee and Meeyoung Cha

*Y'all should read this! Identifying Plurality in Second-Person Personal Pronouns in English Texts*

Gabriel Stanovsky and Ronen Tamari

*An Edit-centric Approach for Wikipedia Article Quality Assessment*

Edison Marrese-Taylor, Pablo Loyola and Yutaka Matsuo

*Additive Compositionality of Word Vectors*

Yeon Seonwoo, Sungjoon Park, Dongkwan Kim and Alice Oh

*Contextualized context2vec*

Kazuki Ashihara, Tomoyuki Kajiwara, Yuki Arase and Satoru Uchida

*Phonetic Normalization for Machine Translation of User Generated Content*

José Carlos Rosales Núñez, Djamé Seddah and Guillaume Wisniewski

**Monday, November, 4, 2019 (continued)**

*Normalization of Indonesian-English Code-Mixed Twitter Data*

Anab Maulana Barik, Rahmad Mahendra and Mirna Adriani

*Unsupervised Neologism Normalization Using Embedding Space Mapping*

Nasser Zalmout, Kapil Thadani and Aasish Pappu

*Lexical Features Are More Vulnerable, Syntactic Features Have More Predictive Power*

Jekaterina Novikova, Aparna Balagopalan, Ksenia Shkaruta and Frank Rudzicz

*Simple Discovery of Aliases from User Comments*

Abram Handler and Brian Clifton

*Towards Actual (Not Operational) Textual Style Transfer Auto-Evaluation*

Richard Yuanzhe Pang

*CodeSwitch-Reddit: Exploration of Written Multilingual Discourse in Online Discussion Forums*

Ella Rabinovich, Masih Sultani and Suzanne Stevenson

**3:00–4:30**      **Poster Session (all papers above)**

**4:30–4:55**      **Coffee Break**

**Monday, November, 4, 2019 (continued)**

**5:00–5:45      Invited Talk: Jing Jiang**

**5:45–6:00      Closing and Best Paper Awards**

