

SocialNLP 2017

**The Fifth International Workshop on
Natural Language Processing for Social Media**

**Proceedings of the Workshop
AFNLP SIG SocialNLP**

EACL 2017 Workshop
April 3, 2017
Valencia, Spain

©2017 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 978-1-945626-42-5

SocialNLP@EACL2017 Chairs' Welcome

It is our great pleasure to welcome you to the Fifth International Workshop on Natural Language Processing for SocialMedia-SocialNLP 2017, associated with EACL 2017. SocialNLP is an interdisciplinary area of natural language processing (NLP) and social computing. We hold SocialNLP twice a year: one in the NLP venue, the other in the associated venue such as those for web technology or artificial intelligence. There are three plausible directions of SocialNLP: (1) addressing issues in social computing using NLP techniques; (2) solving NLP problems using information from social media; and (3) handling new problems related to both social computing and natural language processing. Through this workshop, we anticipate to provide a platform for research outcome presentation and head-to-head discussion in the area of SocialNLP, with the hope to combine the insight and experience of prominent researchers from both NLP and social computing domains to contribute to the area of SocialNLP jointly. The submissions to this year's workshop were again of high quality and we had a competitive selection process. We received 13 submissions from Asia, Europe, and the United States, and due to a rigorous review process, we only accepted 6 long oral papers. Thus the acceptance rate was 46 percent. Compared to the recent workshops in United States, the submission number is a bit small. This also encourages us to have more related activities in Europe to expand our community here.

This year, we are delighted to have Prof. Dirk Hovy, from the University of Copenhagen, as our keynote speaker. We also encourage attendees to attend the keynote and invited talk presentation to have more discussions with outstanding researchers. Their valuable and insightful talk can and will guide us to a better understanding of the future. Putting together SocialNLP 2017 was a team effort. We first thank the authors for providing the quality content of the program. We are grateful to the program committee members, who worked very hard in reviewing papers and providing feedback for authors. Finally, we especially thank the Workshop Committee Chairs Prof. Laura Rimell and Prof. Richard Johansson.

We hope you join our community and enjoy the workshop!

Organizers

Lun-Wei Ku, Academia Sincia, Taiwan

Cheng-Te Li, National Cheng Kung University, Taiwan

Organizers

Lun-Wei Ku, Academia Sincia, Taiwan
Cheng-Te Li, National Cheng Kung University, Taiwan

Program Committee:

Tim Althoff, Stanford University
Sabine Bergler, Concordia University
Berlin Chen, National Taiwan Normal University
Hsin-Hsi Chen, National Taiwan University
Hai Leong Chieu, DSO National Laboratories
Monojit Choudhury, Microsoft Research
Freddy Chua, Singapore Management University
Nigel Collier, University of Cambridge
Danilo Croce, University of Roma, Tor Vergata
Lei Cui, Microsoft Research
Ronan Cummins, University of Cambridge
Pradipto Das, Rakuten USA
Min-Yuh Day, Tamkang University, Taiwan
Ann Devitt, Trinity College Dublin
Eduard Dragut, Temple University
Koji Eguchi, Kobe University
Michael Elhadad, Ben-Gurion University
Wei Gao, Qatar Computing Research Institute
Spandana Gella, University of Edinburgh
Marco Guerini, Fondazione Bruno Kessler
Weiwei Guo, LinkedIn
William Hamilton, Stanford University
Graeme Hirst, University of Toronto
Wen-Lian Hsu, Academia Sinica
Diana Inkpen, University of Ottawa
David Jurgens, Stanford University
Pallika Kanani, Oracle Labs
Soo-Min Kim, Amazon
Roman Klinger, University of Stuttgart
June-Jei Kuo, National Chung Hsing University
Tsung-Ting Kuo, University of California, San Diego
Cheng-Te Li, National Cheng Kung University
Chuan-Jie Lin, National Taiwan Ocean University
Shou-De Lin, National Taiwan University
Yiqun Liu, Tsinghua University
Zhiyuan Liu, Tsinghua University
Bin Lu, Google Inc.
Zhunchen Luo, China Defense Science and Technology Information Center
Bruno Martins, University of Lisbon
Yelena Mejova, Qatar Computing Research Institute
Rada Mihalcea, University of Michigan
Manuel Montes-y-Gómez, INAOE, Mexico

Dong Nguyen, University of Twente
Harris Papageorgiou, ATHENA Research and Innovation Center
Souneil Park, Telefonica Research
Michael Paul, University of Colorado Boulder
Georgios Petasis, NCSR "Demokritos"
Stephen Pulman, Oxford University
Sravana Reddy, Wellesley College
Paolo Rosso, Universitat Politècnica de València
Derek Ruths, McGill University
Saurav Sahay, Intel Labs
Hassan Saif, The Open University
Yohei Seki, University of Tsukuba
Mário J. Silva, Universidade de Lisboa
Yanchuan Sim, Institute for Infocomm Research
Jan Snajder, University of Zagreb
Jannik Strötgen, Max Planck Institute for Informatics
Xavier Tannier, Université Paris-Sud, LIMSI, CNRS
Mike Thelwall, University of Wolverhampton
Ming-Feng Tsai, National Chengchi University
Paola Velardi, Università di Roma
Marc Verhagen, Brandeis University
Svitlana Volkova, PNNL
Xiaojun Wan, Peking University
Hsin-Min Wang, Academia Sinica
Jenq-Haur Wang, National Taipei University of Technology
William Yang Wang, UC Santa Barbara
Ingmar Weber, Qatar Computing Research Institute, HBKU
Robert West, École Polytechnique Fédérale de Lausanne
Kam-Fai Wong, The Chinese University of Hong Kong
Shih-Hung Wu, Chaoyang University of Technology
Ruifeng Xu, Harbin Institute of Technology
Yi Yang, Georgia Tech
Liang-Chih Yu, Yuan Ze University
Zhe Zhang, IBM Watson
Hua-Ping Zhang, Beijing Institute of Technology
Ming Zhou, Microsoft Research Asia
Deyu Zhou, Southeast University

Table of Contents

<i>A Survey on Hate Speech Detection using Natural Language Processing</i> Anna Schmidt and Michael Wiegand.....	1
<i>Facebook sentiment: Reactions and Emojis</i> Ye Tian, Thiago Galery, Giulio Dulcinati, Emilia Molimpakis and Chao Sun.....	11
<i>Potential and Limitations of Cross-Domain Sentiment Classification</i> Jan Milan Deriu, Martin Weilenmann, Dirk Von Gruenigen and Mark Cieliebak.....	17
<i>Aligning Entity Names with Online Aliases on Twitter</i> Kevin McKelvey, Peter Goutzounis, Stephen da Cruz and Nathanael Chambers.....	25
<i>Character-based Neural Embeddings for Tweet Clustering</i> Svitlana Vakulenko, Lyndon Nixon and Mihai Lupu.....	36
<i>A Twitter Corpus and Benchmark Resources for German Sentiment Analysis</i> Mark Cieliebak, Jan Milan Deriu, Dominic Egger and Fatih Uzdilli.....	45

Conference Program

09:30–09:40 Opening

09:40–10:50 *Keynote Speech 1: NLP, the perfect social (media) science?*
Prof. Dirk Hovy, the University of Copenhagen

10:50–11:20 Coffee Break

11:20–12:10 *A Survey on Hate Speech Detection using Natural Language Processing*
Anna Schmidt and Michael Wiegand

12:10–12:40 *Facebook sentiment: Reactions and Emojis*
Ye Tian, Thiago Galery, Giulio Dulcinati, Emilia Molimpakis and Chao Sun

12:40–13:10 *Potential and Limitations of Cross-Domain Sentiment Classification*
Jan Milan Deriu, Martin Weilenmann, Dirk Von Gruenigen and Mark Cieliebak

14:20–15:30 Keynote Speech 2

15:30–16:20 Coffee Break/Poster Session

16:20–16:50 *Aligning Entity Names with Online Aliases on Twitter*
Kevin McKelvey, Peter Goutzounis, Stephen da Cruz and Nathanael Chambers

16:50–17:20 *Character-based Neural Embeddings for Tweet Clustering*
Svitlana Vakulenko, Lyndon Nixon and Mihai Lupu

17:20–17:50 *A Twitter Corpus and Benchmark Resources for German Sentiment Analysis*
Mark Cieliebak, Jan Milan Deriu, Dominic Egger and Fatih Uzdilli

