PEOPLES 2018

Second Workshop on Computational Modeling of PEople's Opinions, PersonaLity, and Emotions in Social media

Proceedings of the Workshop

PEOPLES2018

June 6, 2018
New Orleans, USA

Copyright of each paper stays with the respective authors (or their employers).

Preface

Welcome to the second edition of PEOPLES (Workshop on Computational Modeling of People's Opinions, Personality and Emotions in Social Media), co-located with the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. The first edition was held at the 26th International Conference on Computational Linguistics (COLING 2016) in Osaka, Japan.

The idea of organizing PEOPLES stemmed from two related observations, namely the availability of large amounts of spontaneous data covering a range of personal aspects and the fact that such aspects are usually studied in isolation. Social media users nowadays freely express what is on their mind at any moment in time, at any location, and about virtually anything. These large amounts of spontaneously produced texts open up a unique opportunity to learn more about such users, e.g., predicting demographic variables (age, gender), but also personality types, as well as emotions and opinion expressions. This observation is not new, of course, and this opportunity has largely been exploited in the recent years, with abundant works on sentiment analysis, emotion detection, and personality. However, such traits of human personality and behavior have indeed attracted a substantial amount of attention but have been mostly studied *in isolation*, often in different - but related - communities, such as NLP, CL, AI. Therefore, we thought that the time was ripe to bring these communities a step closer to study people's traits and expressions jointly and in their interplay on such large volumes of available data.

The communities' response, with 25 received submissions coming from 11 different countries and going well beyond typical NLP topics, proves again this year that there is wide interest at this intersection, and we are happy to be able to provide a context for exchanging ideas.

Following the reviewers's advice, 14 papers were selected for inclusion in the proceedings. They cover a wide range of topics related to the three main PEOPLES themes (personality, emotion and opinion), their interaction and the impact of their modeling on social aspects like well-being, political preferences, humor and language use.

To further enrich this volume, we additionally invited our keynote speakers to submit position papers that accompany their talks, and are excited that both of our keynotes submitted excellent papers touching upon issues of making NLP models more demographically aware and how researchers from related fields such as demography can benefit from NLP techniques.

We hope that this is just the second edition of what will become series of workshops bringing together researchers in Computational Linguistics, Natural Language Processing and Computational Social Science, who share an interest in personality, opinion and emotion detection, and especially in researching the intertwining of such traits and expressions.

We would like to thank our program committee consisting of 33 researchers from a variety of backgrounds for their insightful and constructive reviews. Without their support, this workshop would not have been possible. In addition, we thank all authors for submitting papers and making PEOPLES a big success. Also thanks to our two invited speakers, Dirk Hovy and Letizia Mencarini (Bocconi University, Italy), for having accepted to come to the workshop and share their expertise and ideas on PEOPLES' topics. We thank NAACL for hosting us, and in particular the local organizers for their support. Lastly, we are extremely grateful to our sponsors, CELI Language Technologies, and the Computational Linguistics group of the University of Groningen for their financial support, without which this workshop would not have gone through.

We look forward to welcoming you all at PEOPLES 2018 in New Orleans!

Malvina, Viviana, Barbara, Claudia

PEOPLES: https://peopleswksh.github.io/

Organisers

Malvina Nissim, University of Groningen, The Netherlands

Viviana Patti, University of Turin, Italy

Barbara Plank, IT University of Copenhagen, Denmark

Claudia Wagner, Claudia Wagner, University of Koblenz and GESIS - Köln

Programme Committee

Nikolaos Aletras, Sheffield University, UK
Pierpaolo Basile, University of Bari, Italy
Valerio Basile, INRIA Sophia Antipolis Méditerranée, France
Arnim Bleier, GESIS Leibniz Institute for the Social Sciences, Germany
Gosse Bouma, University of Groningen, The Netherlands
Erik Cambria, Nanyang Technological University, Singapore
Fabio Celli, University of Trento, Italy
Chloé Clavel, LTCI-CNRS, Telecom-ParisTech, France
Franco Cutugno, University of Naples Federico II, Italy
Walter Daelemans, University of Antwerp, Belgium
David Garcia, Complexity Science Hub Vienna and Medical University of Vienna, Austria
Ancsa Hannak, Central European University, Hungary
Dan Hardt, Copenhagen Business School, Denmark
Dirk Hovy, Bocconi University, Italy
Richard Johansson, University of Gothenburg, Sweden
David Jurgens, Stanford University, US
Svetlana Kiritchenko, NRC-Canada, Canada
Florian Kuhnemann, Radboud Universiteit Nijmegen, The Netherlands
Fei Liu, Melbourne University, Australia
Nikola Ljubešić, Jožef Stefan Institute, Slovenia
Kim Luyckx, Biomina Research Group, Belgium
Eric Malmi, Aalto University, Finland
Héctor Martínez Alonso, Thomson Reuters, CA
Rada Mihalcea, University of Michigan, US
Saif Mohammad, NRC-Canada, Canada
Dong Nguyen, University of Twente, The Netherlands
Scott Nowson, Accenture Centre for Innovation, Dublin, Ireland
Massimo Poesio, Queen Mary University, UK
Martin Potthast, Leipzig University, Germany
Daniel Preotiuc-Pietro, University of Pennsylvania, US
Paolo Rosso, Technical University of Valencia, Spain
Hassan Saif, Knowledge Media Institute, UK
Ingmar Weber, QCRI, Qatar

Sponsors

Keynote Speakers

The social and the neural network: How to make Natural Language Processing about people again
Dirk Hovy
Bocconi University, Italy

The potential of the computational linguistic analysis of social media for population studies
Letizia Mencarini
Bocconi University, Italy

Table of Contents

Conference Program

Thursday, June 6, 2018

8:50–9:00   *Opening Remarks*

Session 1

9:00–09:20   *What makes us laugh? Investigations into Automatic Humor Classification*
Vikram Ahuja, Taradheesh Bali and Navjyoti Singh

9:20–09:40   *Social and Emotional Correlates of Capitalization on Twitter*
Sophia Chan and Alona Fyshe

9:40–10:00   *Building an annotated dataset of app store reviews with Appraisal features in English and Spanish*
Natalia Mora and Julia Lavid-López

10:00–10:15   *Enabling Deep Learning of Emotion With First-Person Seed Expressions*
Hassan Alhuzali, Muhammad Abdul-Mageed and Lyle Ungar

10:15–10:30   *A Dataset of Hindi-English Code-Mixed Social Media Text for Hate Speech Detection*
Aditya Bohra, Deepanshu Vijay, Vinay Singh, Syed Sarfaraz Akhtar and Manish Shrivastava

10:30–11:00   *Coffee*

Session 2

11:00–12:00   **Keynote:** *The Social and the Neural Network: How to Make Natural Language Processing about People again*
Dirk Hovy

12:00–12:15   *Observational Comparison of Geo-tagged and Randomly-drawn Tweets*
Tom Lippincott and Annabelle Carrell

12:15–12:30   *Johns Hopkins or johnny-hopkins: Classifying Individuals versus Organizations on Twitter*
Zach Wood-Doughty, Praateek Mahajan and Mark Dredze

12:30–14:00   *Lunch*

Thursday, June 6, 2018 (continued)

Session 3

14:00–15:00 **Keynote:** *The Potential of the Computational Linguistic Analysis of Social Media for Population Studies*
Letizia Mencarini

15:00–15:15 *Understanding the Effect of Gender and Stance in Opinion Expression in Debates on "Abortion"*
Esin Durmus and Claire Cardie

15:15–15:30 *Frustrated, Polite, or Formal: Quantifying Feelings and Tone in Email*
Niyati Chhaya, Kushal Chawla, Tanya Goyal, Projjal Chanda and Jaya Singh

15:30–16:00 *Coffee*

Session 4

16:00–16:20 *Reddit: A Gold Mine for Personality Prediction*
Matej Gjurković and Jan Šnajder

16:20–16:40 *Predicting Authorship and Author Traits from Keystroke Dynamics*
Barbara Plank

16:40–17:00 *Predicting Twitter User Demographics from Names Alone*
Zach Wood-Doughty, Nicholas Andrews, Rebecca Marvin and Mark Dredze

17:00–17:15 *Modeling Personality Traits of Filipino Twitter Users*
Edward Tighe and Charibeth Cheng

17:15–17:30 *Grounding the Semantics of Part-of-Day Nouns Worldwide using Twitter*
David Vilares and Carlos Gómez-Rodríguez

17:30–18:00 *Discussion and Closing*