

PEOPLES 2020

**Computational Modeling of PEople's Opinions, PersonaLity,
and Emotions in Social media**

Proceedings of the Third Workshop



December 13, 2020
Barcelona, Spain (Online)

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

ISBN 978-1-952148-49-1

Preface

Welcome to the third edition of PEOPLES, the Workshop on Computational Modeling of People’s Opinions, Personality and Emotions in Social Media, co-located with the 28th International Conference on Computational Linguistics (COLING 2020). PEOPLES 2020 is the continuation of a successful series of workshops that were held at the 26th International Conference on Computational Linguistics (COLING 2016) in Osaka, Japan, and at the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2018), New Orleans, USA, aimed at providing a forum for researchers who share an interest in personality, opinion and emotion detection, as well as the impact of work in this research field on society.

The communities’ response, with 32 received submissions coming from 15 different countries from Asia-Pacific, Europe and the USA, and going well beyond typical Natural Language Processing topics, proves again this year that there is wide interest at the intersection of fields studying sentiment analysis, emotion detection, and personality in disciplines such Computational Linguistics, Natural Language Processing, and Computational Social Science, and we are happy to be able to provide a context for exchanging ideas. Also, we deem it important to provide a forum where crucial discussion about ethical aspects related to the research involving the PEOPLES traits can find a place to be, and to grow.

Following the reviewers’s advice, 15 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 modelling on social aspects like well-being, inclusion, political preferences and language use. This year, we also noticed that multilinguality is definitely a key issue, with a large variety of languages represented in the research described. Additionally, we had specifically mentioned in the call that we would very much welcome research focusing on how the usual PEOPLES dimensions might have changed or might be changing due to the special COVID-related circumstances. Clearly this is a dominant topic in the community from several viewpoints, with at least four papers specifically discussing pandemic-related research.

We are extremely grateful to David Jurgens from the University of Michigan for giving a keynote talk. His work at the interesection of NLP and Computational Social Science enriches the workshop program.

We would like to thank our program committee consisting of 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. A big thanks to our publicity and publication chair Esin Durmus, Cornell University, for her endless support and for compiling these proceedings, particularly in these difficult pandemic times.

We thank COLING for hosting us again, this time virtually, and in particular the local organizers for their support. Thanks for all the extra efforts of moving the conference and all workshops online, unlike the original plans. We hope you and your family are all safe. 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.

We look forward to welcoming you all at this fully virtual edition of PEOPLES 2020!

Malvina, Viviana, Barbara, and Esin

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

Esin Durmus, Cornell University, US

Programme Committee

Jeremy Barnes, University of Oslo, Norway

Maria Barrett, IT University of Copenhagen, Denmark

Pierpaolo Basile, University of Bari, Italy

Valerio Basile, University of Turin, Italy

Elisa Bassignana, University of Turin, Italy

Meriem Beloucif, Universität Hamburg, Germany

Laura Bostan, University of Stuttgart, Germany

Erik Cambria, Nanyang Technological University, Singapore

Tommaso Caselli, University of Groningen, The Netherlands

Chloé Clavel, LTCI-CNRS, Telecom-ParisTech, France

Walter Daelemans, University of Antwerp, Belgium

Dipankar Das, Jadavpur University, India

Esin Durmus, Cornell University, US

Komal Florio, University of Torino, Italy

Tommaso Fornaciari, Bocconi University, Italy

Aparna Garimella, University of Michigan, USA

Albert Gatt, University of Malta, Malta

Rob van de Goot, IT University of Copenhagen, Denmark

Mareike Hartmann, University of Copenhagen, Denmark

Delia Irazú Hernández Farías, Universidad de Guanajuato Campus León, Mexico

Aditya Joshi, CSIRO, AU

Svetlana Kiritchenko, NRC-Canada, Canada

Roman Klinger, University of Stuttgart, Germany

Florian Kunneman, Vrije Universiteit Amsterdam, The Netherlands

Fei Liu, Melbourne University, Australia
Nikola Ljubešić, Jožef Stefan Institute, Slovenia
Kim Luyckx, Antwerp University Hospital, Belgium
Gosse Minema, University of Groningen, Netherlands
Dong Nguyen, Utrecht University, The Netherlands
Malvina Nissim, University of Groningen, The Netherlands
Scott Nowson, Artificial Intelligence Lead, PwC Middle East, United Arab Emirates
Debora Nozza , Bocconi University, Italy
Endang Wahyu Pamungkas, University of Turin, Italy
Viviana Patti, University of Turin, Italy
Paola Pietrandrea, Université de Lille, France
Barbara Plank, IT University of Copenhagen, Denmark
Daniel Preotiuc-Pietro, Bloomberg, USA
Paolo Rosso, Technical University of Valencia, Spain
Ingmar Weber, QCRI, Qatar
Charlie Welch, University of Michigan, USA

Sponsors

PEOPLES 2020 is organized with the support of CELI Language Technology (<https://www.celi.it/en/>) and the Computational Linguistics group of CLCG (<http://www.rug.nl/research/clcg/>), University of Groningen.



university of
 groningen

Keynote Speakers

Social Interpretations of Interpersonal Communication

David Jurgens

University of Michigan, US

Table of Contents

<i>Inferring Neuroticism of Twitter Users by Utilizing their Following Interests</i> Joran Cornelisse	1
<i>Matching Theory and Data with Personal-ITY: What a Corpus of Italian YouTube Comments Reveals About Personality</i> Elisa Bassignana, Malvina Nissim and Viviana Patti	11
<i>Red Is Open-Minded, Blue Is Conscientious: Predicting User Traits From Instagram Image Data</i> Lisa Branz, Patricia Brockmann and Annika Hinze	23
<i>Persuasiveness of News Editorials depending on Ideology and Personality</i> Roxanne El Baff, Khalid Al Khatib, Benno Stein and Henning Wachsmuth	29
<i>HopeEDI: A Multilingual Hope Speech Detection Dataset for Equality, Diversity, and Inclusion</i> Bharathi Raja Chakravarthi	41
<i>KanCMD: Kannada CodeMixed Dataset for Sentiment Analysis and Offensive Language Detection</i> Adeep Hande, Ruba Priyadharshini and Bharathi Raja Chakravarthi	54
<i>Contextual Augmentation of Pretrained Language Models for Emotion Recognition in Conversations</i> Jonggu Kim, Hyeonmok Ko, seoha song, Saebom Jang and Jiyeon Hong	64
<i>Social Media Unrest Prediction during the COVID-19 Pandemic: Neural Implicit Motive Pattern Recognition as Psychometric Signs of Severe Crises</i> Dirk Johannßen and Chris Biemann	74
<i>Topic and Emotion Development among Dutch COVID-19 Twitter Communities in the early Pandemic</i> Boris Marinov, Jennifer Spenader and Tommaso Caselli	87
<i>Sentiments in Russian Medical Professional Discourse during the Covid-19 Pandemic</i> Irina Ovchinnikova, Liana Ermakova and Diana Nurbakova	99
<i>Multilingual Emoticon Prediction of Tweets about COVID-19</i> Stefanos Stoikos and Mike Izbicki	109
<i>Experiencers, Stimuli, or Targets: Which Semantic Roles Enable Machine Learning to Infer the Emotions?</i> Laura Ana Maria Oberländer, Kevin Reich and Roman Klinger	119
<i>Learning Emotion from 100 Observations: Unexpected Robustness of Deep Learning under Strong Data Limitations</i> Sven Buechel, João Sedoc, H. Andrew Schwartz and Lyle Ungar	129
<i>Cross-lingual Emotion Intensity Prediction</i> Irean Navas Alejo, Toni Badia and Jeremy Barnes	140
<i>The LiLaH Emotion Lexicon of Croatian, Dutch and Slovene</i> Nikola Ljubešić, Ilija Markov, Darja Fišer and Walter Daelemans	153

Conference Program

Sunday, December 13, 2020

14:00–15:00 Session 1

14:00–14:10 *Opening Remarks*

14:10–14:30 *Inferring Neuroticism of Twitter Users by Utilizing their Following Interests*
Joran Cornelisse

14:30–14:50 *Matching Theory and Data with Personal-ITY: What a Corpus of Italian YouTube Comments Reveals About Personality*
Elisa Bassignana, Malvina Nissim and Viviana Patti

14:50–15:00 *Red Is Open-Minded, Blue Is Conscientious: Predicting User Traits From Instagram Image Data*
Lisa Branz, Patricia Brockmann and Annika Hinze

15:00–16:00 Session 2

15:00–15:20 *Persuasiveness of News Editorials depending on Ideology and Personality*
Roxanne El Baff, Khalid Al Khatib, Benno Stein and Henning Wachsmuth

15:20–15:30 *HopeEDI: A Multilingual Hope Speech Detection Dataset for Equality, Diversity, and Inclusion*
Bharathi Raja Chakravarthi

15:30–15:40 *KanCMD: Kannada CodeMixed Dataset for Sentiment Analysis and Offensive Language Detection*
Adeep Hande, Ruba Priyadharshini and Bharathi Raja Chakravarthi

15:40–15:50 *Contextual Augmentation of Pretrained Language Models for Emotion Recognition in Conversations*
Jonggu Kim, Hyeonmok Ko, seoha song, Saebom Jang and Jiyeon Hong

15:50–16:00 *TWEETEVAL: Unified Benchmark and Comparative Evaluation for Tweet Classification (Findings of EMNLP 2020)*
Francesco Barbieri, Jose Camacho-Collados, Leonardo Neves, Luis Espinosa-Anke

Sunday, December 13, 2020 (continued)

16:00–17:00 Session 3

16:00–16:45 *Keynote: Social Interpretations of Interpersonal Communication*
David Jurgens (University of Michigan, US)

16:45–17:00 Break

17:00–18:00 Session 4

17:00–17:20 *Social Media Unrest Prediction during the COVID-19 Pandemic: Neural Implicit Motive Pattern Recognition as Psychometric Signs of Severe Crises*
Dirk Johannßen and Chris Biemann

17:20–17:30 *Topic and Emotion Development among Dutch COVID-19 Twitter Communities in the early Pandemic*
Boris Marinov, Jennifer Spenader and Tommaso Caselli

17:30–17:50 *Sentiments in Russian Medical Professional Discourse during the Covid-19 Pandemic*
Irina Ovchinnikova, Liana Ermakova and Diana Nurbakova

17:50–18:00 *Multilingual Emoticon Prediction of Tweets about COVID-19*
Stefanos Stoikos and Mike Izbicki

18:00–19:00 Session 5

18:00–18:20 *Experiencers, Stimuli, or Targets: Which Semantic Roles Enable Machine Learning to Infer the Emotions?*
Laura Ana Maria Oberländer, Kevin Reich and Roman Klinger

18:20–18:30 *Learning Emotion from 100 Observations: Unexpected Robustness of Deep Learning under Strong Data Limitations*
Sven Buechel, João Sedoc, H. Andrew Schwartz and Lyle Ungar

18:30–18:50 *Cross-lingual Emotion Intensity Prediction*
Irean Navas Alejo, Toni Badia and Jeremy Barnes

18:50–19:00 *The LiLaH Emotion Lexicon of Croatian, Dutch and Slovene*
Nikola Ljubešić, Ilija Markov, Darja Fišer and Walter Daelemans

Sunday, December 13, 2020 (continued)

19:10–20:00 *Closing Remarks and QA*

