The Workshop on Natural Language Processing and Computational Social Science

Proceedings of the Third Workshop

June 6, 2019
Minneapolis, USA
Welcome to the Third Workshop on NLP and Computational Social Science!

This workshop series builds on a successful string of iterations, with dozens of interdisciplinary submissions to make NLP techniques and insights standard practice in CSS research. Our focus is on NLP for social sciences - to continue the progress of CSS, and to integrate CSS with current trends and techniques in NLP.

We received 36 submissions, and due to a rigorous review process by our committee, we accepted 12 archival entries, and 2 non-archival abstracts. The program this year includes 6 papers presented as spotlight talks, and 14 posters. We are especially excited to see so many submissions from outside of NLP, and hope to continue the tradition to foster a dialogue between NLP researchers and users of NLP technology in the social sciences. We are also glad to present a fantastic selection of invited speakers from various aspects of computational social science.

We would like to thank all authors of the accepted papers, our invited speakers, and the fantastic organizing committee that made this workshop possible, and, last but not least, all attendees! Special thanks to our sponsors, whose generous contributions have allowed us to support student scholarships and the travel of our invited speakers.

We hope you enjoy the workshop!

The NLP and CSS workshop organizing team.
Organizers :
Svitlana Volkova (Pacific Northwest National Laboratory)
David Jurgens (University of Michigan)
Dirk Hovy (Bocconi University)
David Bamman (UC Berkeley)
Oren Tsur (Ben Gurion University)

Program Committee :

Abram Handler    Ian Stewart    Pierre Nugues
A. Seza Doğruöz  Jacob Eisenstein  Rada Mihalcea
Afshin Rahimi    John Beieker    Rahul Bhagat
Alexandra Balahur Jonathan K. Kummerfeld  Raquel Fernández
Alice Oh         Joseph Hoover    Rebecca Resnik
Amittai Axelrod  Julian Brooke    Reihane Boghrati
Anja Belz        Juri Ganitkevitch  Rob Voigt
Ann Clifton      Kaiping Chen    Roman Klinger
April Foreman    Kenneth Joseph   Sara Rosenthal
Asad Sayeed      Kristen Johnson  Sara Tonelli
Barbara Plank    Kristen M. Altenburger  Sebastian Padó
Brendan O’Connor  Kristy Hollingshead  Seyed Abolghasem Mirroshandel
Caroline Brun    Loring Ingraham  Shachar Mirkin
Carolyn Rose     Ludovic Rheault  Sravana Reddy
Chris Brew       Lyle Ungar      Steve DeNeefe
Christopher Potts Marie-Catherine de Marneffe  Steve Bethard
Courtney Napoles  Marti A. Hearst  Steven Wilson
Dan Goldwasser   Massimo Poesio  Sungiwhan Mac Kim
Dan Simonson     Matko Bošnjak    Swede White
Daniel Préal-Pietro  Michael Bloodgood  Taylor Berg-Kirkpatrick
David Bamman     Michael Heilman    Teresa Lynn
Diana Maynard    Molly Ireland  Thierry Poibeau
Djamé Seddah    Molly Roberts    Valery Dzutsati
Ekaterina Kochmar  Natalie Ahn  Timothy Baldwin
Eric Bell        Natalie Schluter  Vinodkumar Prabhakaran
François Yvon    Nikola Ljubešić  Yvette Graham
Gideon Mann      Oul Han    Zeerak Waseem
Glen Coppersmith  Pedro Rodrigez  
Guenter Neumann  Peter Makarov  
H. Andrew Schwartz  Philipp Koehn  

Invited Speakers:

Lisa Green, University of Massachusetts Amherst
Brent Hecht, Northwestern University
Lana Yarosh, University of Minnesota
Rada Mihalcea, University Of Michigan
# Table of Contents

**Not My President: How Names and Titles Frame Political Figures**  
Esther van den Berg, Katharina Korfhage, Josef Ruppenhofer, Michael Wiegand and Katja Markert  
1

**Identification, Interpretability, and Bayesian Word Embeddings**  
Adam Lauretig .......................................................... 7

**Tweet Classification without the Tweet: An Empirical Examination of User versus Document Attributes**  
Veronica Lynn, Salvatore Giorgi, Niranjan Balasubramanian and H. Andrew Schwartz......... 18

**Geolocating Political Events in Text**  
Andrew Halterman......................................................... 29

**Neural Network Prediction of Censorable Language**  
Kei Yin Ng, Anna Feldman, Jing Peng and Chris Leberknight .................. 40

**Modeling performance differences on cognitive tests using LSTMs and skip-thought vectors trained on reported media consumption.**  
Maury Courtland, Aida Davani, Melissa Reyes, Leigh Yeh, Jun Leung, Brendan Kennedy, Morteza Dehghani and Jason Zevin .............................................................. 47

**Using time series and natural language processing to identify viral moments in the 2016 U.S. Presidential Debate**  
Josephine Lukito, Prathusha K Sarma, Jordan Foley and Aman Abhishek ................. 54

**Stance Classification, Outcome Prediction, and Impact Assessment: NLP Tasks for Studying Group Decision-Making**  
Elijah Mayfield and Alan Black ........................................ 65

**A Sociolinguistic Study of Online Echo Chambers on Twitter**  
Nikita Duseja and Harsh Jhamtani .................................. 78

**Uphill from here: Sentiment patterns in videos from left- and right-wing YouTube news channels**  
Felix Soldner, Justin Chun-ting Ho, Mykola Makhortyk, Isabelle W.J. van der Vegt, Maximilian Mozes and Bennett Kleinberg ......................................................... 84

**Simple dynamic word embeddings for mapping perceptions in the public sphere**  
Nabeel Gillani and Roger Levy ........................................ 94

**Modeling Behavioral Aspects of Social Media Discourse for Moral Classification**  
Kristen Johnson and Dan Goldwasser ................................ 100
Workshop Program

Thursday, June 6, 2019

9:00–10:30  Session 1

9:00–9:45  Invited Talk 1: Using Computational Social Science to Understand the "Social Science of Computing"
Brent Hecht

9:45–10:30  Invited Talk 2: From Words To People And Back Again
Rada Mihalcea

10:30–11:00  Coffee Break

11:00–12:30  Session 2: Poster Session

11:00–12:30  Not My President: How Names and Titles Frame Political Figures
Esther van den Berg, Katharina Korfhage, Josef Ruppenhofer, Michael Wiegand and Katja Markert

11:00–12:30  Identification, Interpretability, and Bayesian Word Embeddings
Adam Lauretig

11:00–12:30  Tweet Classification without the Tweet: An Empirical Examination of User versus Document Attributes
Veronica Lynn, Salvatore Giorgi, Niranjan Balasubramanian and H. Andrew Schwartz

11:00–12:30  Geolocating Political Events in Text
Andrew Halterman

11:00–12:30  Neural Network Prediction of Censorable Language
Kei Yin Ng, Anna Feldman, Jing Peng and Chris Leberknight

11:00–12:30  Modeling performance differences on cognitive tests using LSTMs and skip-thought vectors trained on reported media consumption.
Maury Courtland, Aida Davani, Melissa Reyes, Leigh Yeh, Jun Leung, Brendan Kennedy, Morteza Dehghani and Jason Zevin

12:30–14:00  Lunch break
14:00–15:30  **Session 3: Spotlight Paper Session**

14:00–14:15  *Using time series and natural language processing to identify viral moments in the 2016 U.S. Presidential Debate*
Josephine Lukito, Prathusha K Sarma, Jordan Foley and Aman Abhishek

14:15–14:30  *Stance Classification, Outcome Prediction, and Impact Assessment: NLP Tasks for Studying Group Decision-Making*
Elijah Mayfield and Alan Black

14:30–14:45  *A Sociolinguistic Study of Online Echo Chambers on Twitter*
Nikita Duseja and Harsh Jhamtani

14:45–15:00  *Uphill from here: Sentiment patterns in videos from left- and right-wing YouTube news channels*
Felix Soldner, Justin Chun-ting Ho, Mykola Makhortykh, Isabelle W.J. van der Vegt, Maximilian Mozes and Bennett Kleinberg

15:00–15:15  *Simple dynamic word embeddings for mapping perceptions in the public sphere*
Nabeel Gillani and Roger Levy

15:15–15:30  *Modeling Behavioral Aspects of Social Media Discourse for Moral Classification*
Kristen Johnson and Dan Goldwasser

15:30–16:00  **Afternoon coffee break**

16:00–17:45  **Session 4**

16:00–16:45  *Invited Talk 3: Subtle Differences in Other Varieties of English: Implications for language-related research and technology*
Lisa Green

16:45–17:15  *Invited Talk 4: Treasure Trove or Pandora’s Box? Investigating Unstructured User-Generated Data from Online Support Communities*
Lana Yarosh

17:15–17:30  **Closing remarks and wrap-up**
Organizers