Workshop on Natural Language Processing and Computational Social Science (NLP + CSS 2016)

Held at the Conference on Empirical Methods in Natural Language Processing (EMNLP 2016)

Austin, Texas, USA
5 November 2016

ISBN: 978-1-5108-3148-3

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2016) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2016)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

Relating semantic similarity and semantic association to how humans label other people
Kenneth Joseph and Kathleen M. Carley 1

Identifying News from Tweets
Jesse Freitas and Heng Ji 11

Obfuscating Gender in Social Media Writing
Sravana Reddy and Kevin Knight 17

Social Proof: The Impact of Author Traits on Influence Detection
Sara Rosenthal and Kathy McKeown 27

Generating Politically-Relevant Event Data
John Beieler 37

User profiling with geo-located posts and demographic data
Adam Poulston, Mark Stevenson and Kalina Bontcheva 43

Gov2Vec: Learning Distributed Representations of Institutions and Their Legal Text
John J. Nay 49

#WhoAmI in 160 Characters? Classifying Social Identities Based on Twitter Profile Descriptions
Anna Priante, Djoerd Hiemstra, Tijs van den Broek, Aaqib Saeed, Michel Ehrenhard and Ariana Need 55

Identifying Stance by Analyzing Political Discourse on Twitter
Kristen Johnson and Dan Goldwasser 66

Learning Linguistic Descriptors of User Roles in Online Communities
Alex Wang, William L. Hamilton and Jure Leskovec 76

The Effects of Data Collection Methods in Twitter
Sunghwan Mac Kim, Stephen Wan, Cecile Paris, Jin Brian and Bella Robinson 86

Expressions of Anxiety in Political Texts
Ludovic Rheault 92

Constructing an Annotated Corpus for Protest Event Mining
Peter Makarov, Jasmine Lorenzini and Hanspeter Kriesi 102

Demographer: Extremely Simple Name Demographics
Rebecca Knowles, Josh Carroll and Mark Dredze 108

Bag of What? Simple Noun Phrase Extraction for Text Analysis
Abram Handler, Matthew Denny, Hanna Wallach and Brendan O'Connor 114

News Sentiment and Cross-Country Fluctuations
Samuel Fraiberger 125

The Clinical Panel: *Leveraging Psychological Expertise During NLP Research*
Glen Coppersmith, Kristy Hollingshead, H. Andrew Schwartz, Molly Ireland, Rebecca Resnik, Kate Loveys, April Foreman and Loring Ingraham 132

Are You a Racist or Am I Seeing Things? Annotator Influence on Hate Speech Detection on Twitter
Zeerak Waseem 138

Disentangling Topic Models: A Cross-cultural Analysis of Personal Values through Words
Steven Wilson, Rada Mihalcea, Ryan Boyd and James Pennebaker 143