

NAACL HLT 2018

**Computational Linguistics and Clinical Psychology:
From Keyboard to Clinic**

Proceedings of the Fifth Workshop

©2018 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 978-1-948087-12-4

Introduction

Mental disorders are among one of the most significant global health problems we face, affecting approximately 450 million people worldwide. This total comprises 300 million people with depression, 60 million people with bipolar affective disorder, 23 million people with schizophrenia and other psychoses, amongst the millions of people affected by other mental disorders worldwide (World Health Organization, 2017). The reach and burden of mental disorders is considerable and continues to grow, and mental disorders have a significant detrimental impact on health and functioning, accounting for 32.4% of years lived with disability and 13% of disability-adjusted life years (Vigo, Thornicroft, & Atun, 2016). Moreover, they create substantial economic consequences for all countries: mental disorders cost US\$2.5 trillion globally, and economic output loss due to mental disorders is anticipated to be US\$16.3 trillion worldwide between 2011 and 2030 (Trautmann, Rehm, & Wittchen, 2016). Compared to other global health problems, mental disorders are widespread with significant, long-term disabilities and economic costs associated.

While there are effective prevention strategies and treatments for mental disorders, many of those at risk or affected do not have access. According to Dr. Shekhar Saxena, the Director of Mental Health and Substance Abuse at the World Health Organization, no countries are developed when it comes to mental health. Approximately 75% of those affected by mental disorders do not receive treatment in high-income countries. Worse still, in middle and low-income countries, 89% and 96% of affected individuals, respectively, do not receive treatment. Key barriers to effective treatment include a shortage in supply of trained mental health workers relative to demand for services, and low funding for treatment and prevention. One solution is research and innovation to increase supply of assessments and treatment for mental disorders.

Language technology can support mental health clinicians, service organizations, and individuals with lived experience in many ways. Conversations have traditionally been a fundamental part of the diagnostic and treatment process for mental disorders. A client's language content helps clinicians deduce a diagnosis or monitor treatment effectiveness. Language provides crucial insights into health and functioning, and language data can be found in and outside of treatment contexts in both text and oral form. Applying language technology to mental health opens the door to creating scalable, inexpensive screening measures or risk assessments that may be administered by a wider variety of healthcare professionals in a broad range of contexts. Language technology may also assist with provision of therapy exercises or emotional support beyond treatment settings through tools such as conversational agents. Natural language processing has been used to track community mental health from public discussions in places like Twitter, thus another application may be in public health monitoring, particularly following crisis events in a community. Language analysis has also supported mental health service organizations by triaging posts delivered to crisis workers by degree of suicidal severity, to ensure those in urgent need of care are attended to quickly. Language technology shows incredible promise for assisting the mental health field in more ways than one.

The continuing goal of the CLPsych workshop series is to bring together computational linguistics researchers with clinicians to talk about the ways that language technology can improve mental health. We aim to continue to foster these discussions while building momentum towards the release of tools and data that can be used by mental health clinicians, service organizations, or those with lived experience of a mental health diagnosis. With this in mind, CLPsych strives to communicate relevant computational methods and results clearly to an interdisciplinary audience, and continually tie the work back to its clinical relevance.

The Computational Linguistics and Clinical Psychology (CLPsych) workshop series began at the 2014 annual meeting of the Association of Computational Linguistics (ACL). The first CLPsych workshop

helped to define state of the art language technologies for mental health. Lively discussions were had on the advantages and disadvantages of language tools for mental health, and the workshop's unique clinically-oriented structure was introduced to the ACL community. This unique structure involves including mental health clinicians as discussants to provide real-world insights into potential applications, strengths, and weaknesses of language technologies presented at the workshop. In subsequent years, workshop participation and attendance has continued to grow as more technologists and clinicians have joined the community. Workshops two and three were held at the North American Association for Computational Linguistics and Human Language Technology's (NAACL-HLT) annual meetings in 2015 and 2016, drawing a near doubling of attendance. 2015's workshop also introduced the Shared Task tradition, which, under guidance of Dr. David Milne in subsequent years, pulled together global teams to create message severity triage systems for youth mental health support service, ReachOut.com. 2017's workshop was held at the Association for Computational Linguistics' (ACL) annual meeting in Vancouver, Canada, where the community further increased in size and discussions about the readiness of language technologies for clinical implementation began.

The Fifth Workshop on Computational Linguistics and Clinical Psychology (CLPsych 2018) was held at the North American Association for Computational Linguistics and Human Language Technology's (NAACL-HLT) annual meeting in New Orleans, LA on June 5th. The theme of 2018's workshop was clinical implementation, with the goal of fostering discussion about whether language technologies for mental health are ready to deploy in the clinical world, and what that deployment could look like. Continuing CLPsych's traditional interdisciplinary approach, practicing clinicians and clinical researchers were included as part of our program committee, and were invited to submit papers and serve as discussants of presented work.

2018's workshop had two submission formats: full papers and dataset papers, the latter of which was a new format to the workshop which allowed researchers to describe new, or newly available, datasets that may be of value to the workshop's community. Overall, 23 submissions were received. Accepted submissions were 13 full papers and 1 dataset paper, which were presented as 5 full talks, 5 mini talks, and 4 posters.

2018's workshop hosted a Shared Task competition which was focused on predicting current and future psychological health from childhood essays using longitudinal language and clinical data from the National Child Development Study, also known as the 1958 British Birth Cohort Study. Teams could participate in one of two tasks, which included predicting childhood psychological health and predicting psychological health at age 50, plus an exploratory task on predicting language or frequency of psychological words at age 50 from childhood language and socio-demographics. 20 teams registered and 7 submissions were received in total. Accepted submissions were presented as an additional 5 posters and 1 full talk at the workshop. 2018's shared task was organized by H. Andrew Schwartz, Alissa Goodman, Veronica Lynn, Kate Niederhoffer, Kate Loveys, and Philip Resnik.

We wish to thank all who contributed to the success of CLPsych 2018. This includes all those who submitted papers or participated in the shared task for their fantastic contributions, those who served as members of the Program Committee for their thoughtful reviews, our clinical discussants for their invaluable insights to the clinical utility of language technologies presented, and our shared task organizers for piecing together a novel research task with applications to early intervention. We also wish to thank our generous workshop sponsors, the University of Maryland Center for Health-Related Informatics and Bioimaging (CHIB) and 7 Cups of Tea, as well as the North American chapter of the Association for Computational Linguistics for making this workshop possible.

Kate Loveys, Kate Niederhoffer, Emily Prud'hommeaux, Rebecca Resnik, & Philip Resnik

Organizers:

Kate Niederhoffer, Circadia Labs
Kate Loveys, Qntfy
Rebecca Resnik, Rebecca Resnik & Associates, LLC
Philip Resnik, University of Maryland
Emily Prud'hommeaux, Boston College

Shared Task Organizers:

H. Andrew Schwartz, Stony Brook University
Veronica Lynn, Stony Brook University
Alissa Goodman, University College London
Kate Niederhoffer, Circadia Labs
Kate Loveys, Qntfy
Philip Resnik, University of Maryland

Program Committee:

Glen Coppersmith, Qntfy
Mark Dredze, Johns Hopkins University
Nazli Goharian, Georgetown University
Brian Roark, Google
Kristy Hollingshead, IHMC
Lyle Ungar, University of Pennsylvania
H. Andrew Schwartz, Stony Brook University & University of Pennsylvania
Daniel Preotiuc-Pietro, University of Pennsylvania
Maarten Sap, University of Washington
Rohan Kshirsagar, Columbia University & Koko
Paul Thompson, Dartmouth College
April Foreman, Department of Veteran's Affairs
Loring Ingraham, George Washington University
Shervin Malmasi, Harvard Medical School
William Jarrold, Nuance Communications
Ayah Zirikly, National Institutes of Health (NIH) & Stanford University
Dirk Hovy, University of Copenhagen
Ted Pedersen, University of Minnesota Duluth
Richard Sproat, Google
Craig Bryan, University of Utah
Masoud Rouhizadeh, Johns Hopkins University
J. Ignacio Serrano, Spanish National Research Council (CSIC)
Eric Morley, Goldman Sachs
Mark Rosenstein, Pearson
Patrick Crutchley, Qntfy
Sean Murphy, New York Psychoanalytic Society & Institute
Michael Woodworth, University of British Columbia
Stacey Dershewitz, George Washington University
Jan van Santen, Oregon Health & Science University
Raymond Tucker, Louisiana State University

Graeme Hirst, University of Toronto
Mike Conway, University of Utah
Nan Bernstein Ratner, University of Maryland
Cecilia Ovesdotter Alm, Rochester Institute of Technology
Joseph Costello, Western Michigan University
Frank Rudzicz, University of Toronto
Hiroki Tanaka, Nara Institute of Science and Technology
Antolin Llorente, Penn State University
Tong Liu, Rochester Institute of Technology
Craig Harman, Johns Hopkins University
Alex Fine, Qntfy
Archna Bhatia, IHMC
Kathleen C. Fraser, University of Toronto
Laura Silverman, University of Rochester
Danielle L. Mowery, University of Utah
Jim Sexton, George Washington University
Christopher Homan, Rochester Institute of Technology
Jill Dolata, Oregon Health & Science University
Matthew Purver, Queen Mary University of London
Dimitrios Kokkinakis, University of Gothenburg
Veronica Lynn, Stony Brook University

Table of Contents

<i>What type of happiness are you looking for? - A closer look at detecting mental health from language</i> Alina Arseniev-Koehler, Sharon Mozgai and Stefan Scherer	1
<i>A Linguistically-Informed Fusion Approach for Multimodal Depression Detection</i> Michelle Morales, Stefan Scherer and Rivka Levitan	13
<i>Expert, Crowdsourced, and Machine Assessment of Suicide Risk via Online Postings</i> Han-Chin Shing, Suraj Nair, Ayah Zirikly, Meir Friedenberg, Hal Daumé III and Philip Resnik .	25
<i>CLPsych 2018 Shared Task: Predicting Current and Future Psychological Health from Childhood Essays</i> Veronica Lynn, Alissa Goodman, Kate Niederhoffer, Kate Loveys, Philip Resnik and H. Andrew Schwartz	37
<i>An Approach to the CLPsych 2018 Shared Task Using Top-Down Text Representation and Simple Bottom-Up Model Selection</i> Micah Iserman, Molly Ireland, Andrew Littlefield, Tyler Davis and Sage Maliepaard	47
<i>Using contextual information for automatic triage of posts in a peer-support forum</i> Edgar Altszyler, Ariel J. Berenstein, David Milne, Rafael A. Calvo and Diego Fernandez Slezak	57
<i>Hierarchical neural model with attention mechanisms for the classification of social media text related to mental health</i> Julia Ive, George Gkotsis, Rina Dutta, Robert Stewart and Sumithra Velupillai	69
<i>Cross-cultural differences in language markers of depression online</i> Kate Loveys, Jonathan Torrez, Alex Fine, Glen Moriarty and Glen Coppersmith	78
<i>Deep Learning for Depression Detection of Twitter Users</i> Ahmed Hussein Orabi, Prasadith Buddhitha, Mahmoud Hussein Orabi and Diana Inkpen	88
<i>Current and Future Psychological Health Prediction using Language and Socio-Demographics of Children for the CLPsych 2018 Shared Task</i> Sharath Chandra Guntuku, Salvatore Giorgi and Lyle Ungar	98
<i>Predicting Psychological Health from Childhood Essays with Convolutional Neural Networks for the CLPsych 2018 Shared Task (Team UKNLP)</i> Anthony Rios, Tung Tran and Ramakanth Kavuluru	107
<i>A Psychologically Informed Approach to CLPsych Shared Task 2018</i> Almog Simchon and Michael Gilead	113
<i>Predicting Psychological Health from Childhood Essays. The UGent-IDLab CLPsych 2018 Shared Task System.</i> Klim Zaporjets, Lucas Sterckx, Johannes Deleu, Thomas Demeester and Chris Develder	119
<i>Can adult mental health be predicted by childhood future-self narratives? Insights from the CLPsych 2018 Shared Task</i> Kylie Radford, Louise Lavrencic, Ruth Peters, Kim Kiely, Ben Hachey, Scott Nowson and Will Radford	126
<i>Automatic Detection of Incoherent Speech for Diagnosing Schizophrenia</i> Dan Iter, Jong Yoon and Dan Jurafsky	136

<i>Oral-Motor and Lexical Diversity During Naturalistic Conversations in Adults with Autism Spectrum Disorder</i>	
Julia Parish-Morris, Evangelos Sariyanidi, Casey Zampella, G. Keith Bartley, Emily Ferguson, Ashley A. Pallathra, Leila Bateman, Samantha Plate, Meredith Cola, Juhi Pandey, Edward S. Brodtkin, Robert T. Schultz and Birkan Tunc	147
<i>Dynamics of an idiosyncrasy of a Russian suicidal blogger</i>	
Tatiana Litvinova, Olga Litvinova and Pavel Seredin	158
<i>RSDD-Time: Temporal Annotation of Self-Reported Mental Health Diagnoses</i>	
Sean MacAvaney, Bart Desmet, Arman Cohan, Luca Soldaini, Andrew Yates, Ayah Zirikly and Nazli Goharian	168
<i>Predicting Human Trustfulness from Facebook Language</i>	
Mohammadzaman Zamani, Anneke Buffone and H. Andrew Schwartz	174
<i>Within and Between-Person Differences in Language Used Across Anxiety Support and Neutral Reddit Communities</i>	
Molly Ireland and Micah Iserman	182
<i>Helping or Hurting? Predicting Changes in Users' Risk of Self-Harm Through Online Community Interactions</i>	
Luca Soldaini, Timothy Walsh, Arman Cohan, Julien Han and Nazli Goharian	194

Conference Program

Tuesday June 5, 2018

9:00–9:15 Opening Remarks

9:15–10:35 Workshop Session I: Presentations with Discussant Commentary

What type of happiness are you looking for? - A closer look at detecting mental health from language

Alina Arseniev-Koehler, Sharon Mozgai and Stefan Scherer

A Linguistically-Informed Fusion Approach for Multimodal Depression Detection

Michelle Morales, Stefan Scherer and Rivka Levitan

Expert, Crowdsourced, and Machine Assessment of Suicide Risk via Online Postings

Han-Chin Shing, Suraj Nair, Ayah Zirikly, Meir Friedenberg, Hal Daumé III and Philip Resnik

10:35–10:55 Break

10:55–11:40 Plenary Session

11:40–12:40 Workshop Session II: Shared Task Presentations with Discussant Commentary

CLPsych 2018 Shared Task: Predicting Current and Future Psychological Health from Childhood Essays

Veronica Lynn, Alissa Goodman, Kate Niederhoffer, Kate Loveys, Philip Resnik and H. Andrew Schwartz

An Approach to the CLPsych 2018 Shared Task Using Top-Down Text Representation and Simple Bottom-Up Model Selection

Micah Iserman, Molly Ireland, Andrew Littlefield, Tyler Davis and Sage Maliepaard

Tuesday June 5, 2018 (continued)

12:40–14:00 Lunch and Poster Session

Using contextual information for automatic triage of posts in a peer-support forum

Edgar Altszyler, Ariel J. Berenstein, David Milne, Rafael A. Calvo and Diego Fernandez Slezak

Hierarchical neural model with attention mechanisms for the classification of social media text related to mental health

Julia Ive, George Gkotsis, Rina Dutta, Robert Stewart and Sumithra Velupillai

Cross-cultural differences in language markers of depression online

Kate Loveys, Jonathan Torrez, Alex Fine, Glen Moriarty and Glen Coppersmith

Deep Learning for Depression Detection of Twitter Users

Ahmed Husseini Orabi, Prasadith Buddhitha, Mahmoud Husseini Orabi and Diana Inkpen

Current and Future Psychological Health Prediction using Language and Socio-Demographics of Children for the CLPsych 2018 Shared Task

Sharath Chandra Guntuku, Salvatore Giorgi and Lyle Ungar

Predicting Psychological Health from Childhood Essays with Convolutional Neural Networks for the CLPsych 2018 Shared Task (Team UKNLP)

Anthony Rios, Tung Tran and Ramakanth Kavuluru

A Psychologically Informed Approach to CLPsych Shared Task 2018

Almog Simchon and Michael Gilead

Predicting Psychological Health from Childhood Essays. The UGent-IDLab CLPsych 2018 Shared Task System.

Klim Zaporjets, Lucas Sterckx, Johannes Deleu, Thomas Demeester and Chris Develder

Can adult mental health be predicted by childhood future-self narratives? Insights from the CLPsych 2018 Shared Task

Kylie Radford, Louise Lavrencic, Ruth Peters, Kim Kiely, Ben Hachey, Scott Nowson and Will Radford

Tuesday June 5, 2018 (continued)

14:00–15:00 Workshop Session III: Presentations with Discussant Commentary

Automatic Detection of Incoherent Speech for Diagnosing Schizophrenia

Dan Iter, Jong Yoon and Dan Jurafsky

Oral-Motor and Lexical Diversity During Naturalistic Conversations in Adults with Autism Spectrum Disorder

Julia Parish-Morris, Evangelos Sariyanidi, Casey Zampella, G. Keith Bartley, Emily Ferguson, Ashley A. Pallathra, Leila Bateman, Samantha Plate, Meredith Cola, Juhi Pandey, Edward S. Brodtkin, Robert T. Schultz and Birkan Tunc

15:00–15:30 Workshop Session IV: Short Presentations

Dynamics of an idiosyncrasy of a Russian suicidal blogger

Tatiana Litvinova, Olga Litvinova and Pavel Seredin

RSDD-Time: Temporal Annotation of Self-Reported Mental Health Diagnoses

Sean MacAvaney, Bart Desmet, Arman Cohan, Luca Soldaini, Andrew Yates, Ayah Zirikly and Nazli Goharian

15:30–16:00 Break

16:00–16:45 Workshop Session V: Short Presentations

Predicting Human Trustfulness from Facebook Language

Mohammadzaman Zamani, Anneke Buffone and H. Andrew Schwartz

Within and Between-Person Differences in Language Used Across Anxiety Support and Neutral Reddit Communities

Molly Ireland and Micah Iserman

Helping or Hurting? Predicting Changes in Users' Risk of Self-Harm Through Online Community Interactions

Luca Soldaini, Timothy Walsh, Arman Cohan, Julien Han and Nazli Goharian

Tuesday June 5, 2018 (continued)

16:45–17:30 Workshop General Discussion and Closing Remarks