Journals   Books   Sign in

Current Opinion in Psychology

Volume 9, June 2016, Pages 77–82

Social media and applications to health behavior

Social media, big data, and mental health: current advances and ethical implications

Mike Conway[1],   , Daniel O'Connor[2]
Show more

http://dx.doi.org/10.1016/j.copsyc.2016.01.004   Get rights and content

Highlights

- Mental health is the fifth greatest contributor to the global burden of disease.
- Population mental health systems require strengthening to address this need.
- Social media big data combined with NLP can address public health research questions.

Mental health (including substance abuse) is the fifth greatest contributor to the global burden of disease, with an economic cost estimated to be US $2.5 trillion in 2010, and expected to double by 2030. Developing information systems to support and strengthen population-level mental health monitoring forms a core part of the World Health Organization's Comprehensive Action Plan 2013–2020. In this paper, we review recent work that utilizes social media 'big data' in conjunction with associated technologies like natural language processing and machine learning to address pressing problems in population-level mental health surveillance and research, focusing both on technological advances and core ethical challenges.

Introduction

Mental illness (including substance abuse) is the fifth greatest contributor to the global burden of disease [1 and 2]. The economic cost of mental illness was estimated to be US $2.5 trillion in 2010, and is expected to double by 2030 [3]. A core goal of the World Health Organization's Comprehensive Mental Health Action Plan 2013–20 is to strengthen information systems for mental health, including increasing capacity for population health monitoring [4]. The widespread use of social media combined with the rapid development of computational infrastructures to support efficient processing of 'big data',[1] and crucially, the maturation of Natural Language Processing (NLP) and Machine

source in the political [6], business [7], and policy [8] contexts, is increasingly used in population health monitoring, and is beginning to be used for mental health applications. Social media analysis is particularly promising in the mental health domain, as Twitter, Facebook, etc., provide access to naturalistic, first person accounts of user behavior, thoughts, and feelings that may be indicative of emotional wellbeing.

An important feature of research in this domain is that it is inherently interdisciplinary and dispersed across health journals (e.g. PubMed), psychology journals (e.g. PsycINFO), and computer science conference and workshop proceedings (e.g. Compendex).[2] This review briefly surveys social media-based applications of NLP to the mental health domain, focusing on both recent technological advances and core ethical issues from the perspective of population-level mental health monitoring.[3]

Mining social media for health

The use of social media 'big data' for health applications — particularly public health applications — is a rapidly growing area of research [10 and 11] variously referred to as *infoveillance* [12], *digital epidemiology* [13], and *digital disease detection* [14]. Twitter in particular, due to its public Application Programming Interface,[4] and status as a 'broadcast' social network[5] has been used for population-level influenza surveillance [16, 17 and 18], monitoring mass gatherings [19 and 20], understanding public sentiment towards vaccination [21], building pharmacovigilance applications (e.g. post-market surveillance of adverse drug events) [22 and 23], understanding public attitudes towards new and emerging tobacco products and e-cigarette marketing [24 and 25], and investigating prescription drug abuse [26].

Mental health and natural language processing

Mental health has been a subject of research for NLP researchers since the early days of the discipline, as evidenced by Weizenbaum's ELIZA interactive Rogerian psychotherapist program [27] (1966), and Colby's 'paranoid' conversational agent, PARRY [28] (1972). As is to be expected, the field has moved on significantly since the development of these early chatbots. Recent work uses sophisticated NLP and ML methods to, for instance, assess suicide risk in pediatric populations based on writing samples [29], predict depression severity and optimal treatment based on narrative text derived from Electronic Health Records [30], identify linguistic features characteristic of early stage dementia [31], and predict the suicide risk of active duty military personnel based on Electronic Health Record data [32]. In parallel with these advances in NLP, there is a rich tradition in the psychology domain (exemplified by Pennebaker [33]) of using carefully developed and validated lexicons organized into various categories (e.g. *anxiety*, *insight*, *achievement*) in order to score texts according to the presence or absence of psychological terms.

Social media, natural language processing, and mental health

Social media has been used extensively in marketing for *sentiment analysis* (broadly, the ascription of positive or negative emotional valence to a text [34]) and for quantifying specific personality traits or dimensions. For example, predicting 'dark triad' traits (i.e. *narcissism*, *Machiavellianism*, and *psychopathy*) from tweets [35], detecting evidence of psychopathy [36], and the identification of 'Big 5' personality dimensions from Facebook data [37**]. Specifically focused on mental health, negative-emotion language on Twitter has been shown to correlate well with official United States suicide statistics at the state level [38**].

De Choudhury

With colleagues at Microsoft Research and Georgia Tech, De Choudhury has been responsible for a pioneering series of papers on applying computational methods to the investigation of mental health issues in a number of different social media platforms, including Twitter [39**, 40 and 41], Facebook [42], and Reddit [43 and 44]. De Choudhury's work has focused on developing methods for both *monitoring population health* and *identifying risk factors for individuals*. In the *population health domain*, De Choudhury *et al.* [39**] describe the creation of a crowdsourced data set of tweets derived from Twitter users with depression-indicative CES-D (*Center for Epidemiological Studies-Depression*) scores. This data-set was then used to train a statistical ML algorithm

control depression data in the identifying new factors for individuals certain, De Choudhury et al. [40] investigated new mothers' experiences of postpartum depression by automatically identifying birth announcements from public Twitter data using cue phrases (e.g. 'it's a boy/girl!'), then analyzing characteristics of the new mothers' Twitter stream before and after birth, discovering that using ML techniques in conjunction with an analysis of pre-birth behavior patterns can predict postnatal emotional and behavioral changes with 71% accuracy.

CLPsych conference

The CLPsych — Computational Linguistics and Clinical Psychology — workshop series has provided an important forum for computer science researchers with an interest in clinical psychology, and for research psychologists and mental health clinicians with an interest in technology. While covering a wide range of mental health applications (e.g. automatically coding therapist/patient interactions [45], and automatically quantifying autistic childrens' repetitive linguistic behavior [46]) the workshop has had a specific focus on population mental health and social media. In particular, participants at the workshop introduced a novel method for developing data sets for specific mental illnesses: pulling tweets (via the public Twitter Application Programming Interface) from users with a self-disclosed, publicly-stated psychiatric diagnosis (e.g. 'I was diagnosed with having P.T.S.D', 'she diagnosed me with anxiety and depression'). The approach was first used to generate a data set for post-traumatic stress disorder, depression, bipolar disorder and seasonal affective disorder [47••], before extending the approach to other conditions (attention deficit hyperactivity disorder, anxiety, borderline, eating disorders, obsessive–compulsive disorder, and schizophrenia) [48]. Work has focused on characterizing language associated with particular mental health conditions on Twitter using variety of methods. Mitchell et al. [49] investigated linguistic characteristics associated with those Twitter users who had a self-disclosed schizophrenia diagnosis, discovering that — when compared to community controls — schizophrenia sufferers were more likely to use the first person, and less likely to use emoticons and exclamation marks — findings consistent with current understanding of schizophrenia (i.e. *preoccupation with self* and *flat affect*, respectively). Using the same dataset as [47••], Preoţiuc-Pietro et al. leveraged NLP techniques to examine 'Big-5' personality and demographic characteristics associated with a self-disclosed diagnosis of depression or PTSD [50], finding that PTSD sufferers were both older and more conscientious than depression sufferers. Resnick et al. [51] used a sophisticated topic modeling ML technique to identify themes in the depression Twitter data generated by Coppersmith et al. [47••], and discovered that the process of aggregating tweets — that is not treating individual tweets as atomic, but rather providing more context by processing data derived from a single user in weekly chunks — substantially improved the quality of results. Mowery et al. took a different approach, manually building and refining an annotation scheme (*coding scheme*) and corpus of Twitter data coded using DSM-5 depression criteria (e.g. *diminished ability to think or concentrate*, *anhedonia*) and psychosocial stressors (e.g. *housing problem*, *occupational problem*) [52], with the goal of creating a shared resource for training and testing algorithms to identify depression symptoms from social media data, and training NLP algorithms for estimating population-level prevalence of depression. Schwartz et al. used Facebook status updates, in combination with the results of a personality survey of 28 749 Facebook users to predict — using a regression model — *degree of depression* for a given user, finding that user mood worsens in the transition from summer to winter [53].

World Well Being Project

Located at the University of Pennsylvania and informed by ideas from positive psychology [54], The World Well Being Project (WWBP)[6] is a collaboration between psychologists, computer scientists, and statisticians to study the psychosocial processes related to health and happiness as manifest in the language of social media. In collaboration with colleagues at the University of Cambridge,[7] WWBP researchers used data derived from users of myPersonality, a Facebook app designed to measure personality variables (including 'Big 5' variables). Using a sample consisting of 71 556 participants who had both completed the online personality questionnaire, and granted access to their Facebook status updates, the researchers found fair to good correlations between personality scores and linguistic features [55 and 37••]. Focusing on the

conjunction with United States Centers for Disease Control mortality data to investigate the correlation between words characteristic of negative emotions (e.g. hostility, disengagement) and heart disease mortality at the US county level, discovering that negative emotions in Twitter were highly correlated with heart diseases mortality figures (indeed, more highly correlated than official socio-economic, demographic, and health statistics) [56].

Ethical implications

As the above review shows, social media analysis can provide access to naturalistic first person accounts of user behavior and opinions that may be indicative of mental health status, enabling researchers to make population-level inferences. The use of social media for health research has been shown to have specific ethical implications regarding: (1) users' expectations regarding the distinction between public and private content [57 and 58], (2) user privacy [59 and 60], (3) and researcher responsibilities [61 and 62••]. All of these pertain to the particular kinds of social media research outlined in the above review.

User expectations

The primary implication of the research detailed in our review is that anything and everything an individual posts to a social media site may be used for research purposes. However, simply because social media is public, and in some cases freely available, it does not follow that it is always ethically appropriate to use it for any research purpose, particularly in relation to sensitive domains such as mental health.

User privacy

Privacy has been identified as a key ethical concern for population-level social media research [61]. Research focused on automatically identifying those who suffer from a given mental illness at the individual, as opposed to population level, can be said to challenge privacy through the association of users with a potentially stigmatizing medical condition. However, the large-scale nature of the data sets in use mean that it is unlikely that individual users will be specifically identified. The potential challenge to privacy occurs here not in the reading or accessing of individual materials (publicly available as they are), but rather in the processing and dissemination of those materials in a way unintended (and potentially even disagreed with) by the users as a group.

Researchers' responsibilities

The expectations and privacy of social media users are salient ethical factors in the research we describe in this review. This does not mean that such research is ethically flawed, especially given the potential benefits of the research at both the individual and population levels. The privacy concerns we raise here focus largely on stigmatization, and place upon researchers the obligation to be sensitive to the scale and generalizability of the conclusions drawn about mental health from social media data.

Conclusion

Recent technological advances hold significant promise for understanding and improving mental health at both the individual and population level. However, risks — particularly to privacy — remain. Researchers should take seriously the notion that the conclusions they draw from these data sets may have very personal, even private implications.
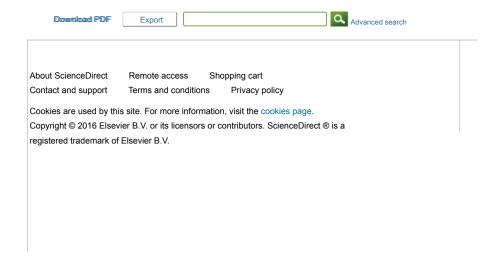
References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- •• of outstanding interest

Acknowledgements

Download PDF   Export   [   ] 🔍 Advanced search

About ScienceDirect   Remote access   Shopping cart
Contact and support   Terms and conditions   Privacy policy