Author Manuscript
Author Manuscript
Author Manuscript

HHS Public Access
Author manuscript
Curr Opin Psychol. Author manuscript; available in PMC 2017 June 01.

Published in final edited form as:
Curr Opin Psychol. 2016 June ; 9: 77–82. doi:10.1016/j.copsyc.2016.01.004.

Social Media, Big Data, and Mental Health: Current Advances and Ethical Implications

Mike Conway[a,*] and **Daniel O'Connor**[b]

[a]Department of Biomedical Informatics, University of Utah, 421 Wakara Way, Salt Lake City, Utah, United States

[b]Wellcome Trust, Gibbs Building, 215 Euston Road, London NW1 2BE, United Kingdom

Abstract

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, 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 Learning (ML) technologies, offers exciting possibilities for the improvement of both population-level and individual-level health. Social media is well established as a data 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

*Corresponding author: mike.conway@utah.edu (Mike Conway).

[1]The term "big data" lacks an agreed definition, but one common formulation characterizes the distinction between "big data" and more traditional data in terms of *velocity*, *volume*, and *variety* (the "three Vs") [5].

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, 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 [16–18], monitoring mass gatherings [19, 20], understanding public sentiment towards vaccination [21], building pharmacovigillance applications (e.g. post-market surveillance of adverse drug events) [22, 23], understanding public attitudes towards new and emerging tobacco products and e-cigarette marketing [24, 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.

[2]Note that in computer science, peer-reviewed conference and workshop papers, as opposed to journals, are the preferred means of disseminating research results.
[3]Note that this review *does not* focus on intervention-based studies (e.g. Facebook's 2014 "emotional contagion" intervention study[9])
[4]Twitter offers several freely accessible Application Programming Interfaces.
[5]Twitter's open status can be contrasted with sources of internet-derived public health data, like *Google Flu Trends* [15] which are not easily accessible by researchers.

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 [39–41], Facebook [42], and Reddit [43, 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]** describes 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 capable of identifying depression-indicative tweets and then applied to geocoded Twitter data derived from 50 US states, with results correlating well with US Centers for Disease Control depression data. In the *identifying risk factors for individuals domain*, 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 [51], 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

Based 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 [37, 55]. Focusing on the population-level impact of psychosocial factors on heart disease mortality, WWBP researchers uses 148 million tweets geocoded at the United States county level in 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].

[6] World Well Being Project: http://wwbp.org
[7] myPersonality Project: http://mypersonality.org

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, 58], (2) user privacy [59, 60], (3) and researcher responsibilities [61, 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.

Acknowledgments

We would like to thank Nicholas Perry (Department of Psychology, University of Utah) and Danielle Mowery (Department of Biomedical Informatics, University of Utah) for their comments on an early draft of this work. Author Mike Conway was partially supported by a grant from the National Library of Medicine (R00LM011393). Daniel O'Connor contributed to this article in a personal capacity; the views expressed are his own and do not necessarily represent the views of the Wellcome Trust.

References

1. Ferrari AJ, Norman RE, Freedman G, Baxter AJ, Pirkis JE, Harris MG, Page A, Carnahan E, Degenhardt L, Vos T, Whiteford HA. The burden attributable to mental and substance use disorders as risk factors for suicide: findings from the global burden of disease study 2010. PLoS One. 2014; 9(4):e91936. [PubMed: 24694747]
2. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, Charlson FJ, Norman RE, Flaxman AD, Johns N, Burstein R, Murray CJL, Vos T. Global burden of disease attributable to mental and substance use disorders: findings from the global burden of disease study 2010. Lancet. 2013; 382(9904):1575–86. [PubMed: 23993280]
3. Bloom, DE.; Cafiero, E.; Jané-Llopis, E.; Abrahams-Gessel, S.; Bloom, LR.; Fathima, S.; Feigl, AB.; Gaziano, T.; Hamandi, A.; Mowafi, M., et al. The global economic burden of noncommunicable diseases, Working paper 87. Program on the Global Demography of Aging. 2012. http://www.hsph.harvard.edu/program-on-the-global-demography-of-aging/WorkingPapers/2012/PGDA_WP_87.pdf
4. World Health Organization. Mental health action plan 2013 – 2020. Tech rep. 2013. http://www.who.int/mental_health/publications/action_plan/en/
5. Lane, J.; Stodden, V.; Bender, S.; Nissenbaum, H. Privacy, big data, and the public good. Cambridge University Press; 2014.
6. Tumasjan A, Sprenger TO, Sandner PG, Welpe IM. Election forecasts with Twitter: How 140 characters reflect the political landscape. Social Science Computer Review. 2011; 29:402–418.
7. Bollen J, Mao H, Zeng X. Twitter mood predicts the stock market. Journal of Computational Science. 2011; 2(1):1–8.
8. Kim D, Kim J. Public opinion sensing and trend analysis on social media: A study on nuclear power on Twitter. International Journal of Multimedia and Ubiquitious Computing. 2014; 9(11):373–384.
9. Kramer ADI, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. Proceedings of the National Academy of Sciences of the United States of America. 2014; 111(24):8788–90. [PubMed: 24889601]
10. Dredze M. How social media will change public health. IEEE Intelligent Systems. 2012; 27(4):81–84.
11. Paul MJ, Dredze M. You are what you tweet: Analyzing Twitter for public health. ICWSM. 2011:265–272.
12. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the internet. J Med Internet Res. 2009; 11(1):e11. [PubMed: 19329408]
13. Salathé M, Bengtsson L, Bodnar TJ, Brewer DD, Brownstein JS, Buckee C, Campbell EM, Cattuto C, Khandelwal S, Mabry PL, Vespignani A. Digital epidemiology. PLoS Comput Biol. 2012; 8(7):e1002616. [PubMed: 22844241]
14. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection–harnessing the web for public health surveillance. N Engl J Med. 2009; 360(21):2153–5. 2157. [PubMed: 19423867]
15. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature. 2009; 457(7232):1012–4. [PubMed: 19020500]
16. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of tweets during the 2009 H1N1 outbreak. PLoS One. 2010; 5(11):e14118. [PubMed: 21124761]
17. Collier N, Son NT, Nguyen NM. OMG U got flu? Analysis of shared health messages for bio-surveillance. J Biomed Semantics. 2011; 2(Suppl 5):S9. [PubMed: 22166368]

18. Broniatowski DA, Paul MJ, Dredze M. National and local influenza surveillance through Twitter: an analysis of the 2012–2013 influenza epidemic. PLoS One. 2013; 8(12):e83672. [PubMed: 24349542]
19. Nsoesie EO, Kluberg SA, Mekaru SR, Majumder MS, Khan K, Hay SI, Brownstein JS. New digital technologies for the surveillance of infectious diseases at mass gathering events. Clin Microbiol Infect. 2015; 21(2):134–40. [PubMed: 25636385]
20. Yom-Tov E, Borsa D, Cox IJ, McKendry RA. Detecting disease outbreaks in mass gatherings using internet data. J Med Internet Res. 2014; 16(6):e154. [PubMed: 24943128]
21. Salathé M, Khandelwal S. Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. PLoS Comput Biol. 2011; 7(10):e1002199. [PubMed: 22022249]
22. Freifeld CC, Brownstein JS, Menone CM, Bao W, Filice R, Kass-Hout T, Dasgupta N. Digital drug safety surveillance: monitoring pharmaceutical products in Twitter. Drug Saf. 2014; 37(5):343–50. [PubMed: 24777653]
23. O'Connor K, Pimpalkhute P, Nikfarjam A, Ginn R, Smith KL, Gonzalez G. Pharmacovigilance on Twitter? mining tweets for adverse drug reactions. AMIA Annu Symp Proc. 2014; 2014:924–33. [PubMed: 25954400]
24. Myslín M, Zhu S-H, Chapman W, Conway M. Using Twitter to examine smoking behavior and perceptions of emerging tobacco products. J Med Internet Res. 2013; 15(8):e174. [PubMed: 23989137]
25. Huang J, Kornfield R, Szczypka G, Emery SL. A cross-sectional examination of marketing of electronic cigarettes on Twitter. Tob Control. 2014; 23(Suppl 3):iii26–30. [PubMed: 24935894]
26. Hanson CL, Burton SH, Giraud-Carrier C, West JH, Barnes MD, Hansen B. Tweaking and tweeting: exploring Twitter for non-medical use of a psychostimulant drug (adderall) among college students. J Med Internet Res. 2013; 15(4):e62. [PubMed: 23594933]
27. Weizenbaum J. Eliza—a computer program for the study of natural language communication between man and machine. Communications of the ACM. 1966; 9(1):36–45.
28. Colby, KM. Artificial Paranoia: A Computer Simulation of Paranoid Processes. Elsevier Science Inc; New York, NY, USA: 1975.
29. Pestian JP, Matykiewicz P, Linn-Gust M, South B, Uzuner O, Wiebe J, Cohen KB, Hurdle J, Brew C. Sentiment analysis of suicide notes: A shared task. Biomed Inform Insights. 2012; 5(Suppl 1): 3–16. [PubMed: 22419877]
30. Huang SH, LePendu P, Iyer SV, Tai-Seale M, Carrell D, Shah NH. Toward personalizing treatment for depression: predicting diagnosis and severity. J Am Med Inform Assoc. 2014; 21(6):1069–75. [PubMed: 24988898]
31. Xuan L, Lancashire I, Hirst G, Jokel R. Longitudinal detection of dementia through lexical and syntactic changes in writing: a case study of three British novelists. Literary and Linguistic Computing. 26:435–461.
32. Poulin C, Shiner B, Thompson P, Vepstas L, Young-Xu Y, Goertzel B, Watts B, Flashman L, McAllister T. Predicting the risk of suicide by analyzing the text of clinical notes. PLoS One. 2014; 9(1):e85733. [PubMed: 24489669]
33. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. Journal of Language and Social Psychology. 2010; 29(1):24–54.
34. Taboada M, Brooke J, Tofiloski M, Voll K, Stede M. Lexicon-based methods for sentiment analysis. Computational Linguistics. 2011; 37(2):267–307.
35. Sumner, C.; Byers, A.; Boochever, R.; Park, GJ. Predicting dark triad personality traits from Twitter usage and a linguistic analysis of tweets. 11th International Conference on Machine Learning and Applications; ICMLA, Boca Raton, FL, USA. December 12–15, 2012; 2012. p. 386-393.http://dx.doi.org/10.1109/ICMLA.2012.218
36. Wald, R.; Khoshgoftaar, TM.; Napolitano, A.; Sumner, C. Using Twitter content to predict psychopathy. 11th International Conference on Machine Learning and Applications; ICMLA, Boca Raton, FL, USA. December 12–15, 2012; 2012. p. 394-401.http://dx.doi.org/10.1109/ICMLA.2012.228

37**. Park G, Schwartz HA, Eichstaedt JC, Kern ML, Kosinski M, Stillwell DJ, Ungar LH, Seligman MEP. Automatic personality assessment through social media language. Journal of Personality and Social Psychology. 2014; 108(6):934–952. Working with Facebook personality questionnaire data derived from the myPersonality Project, researchers at the University of Pennsylvania and the University of Cambridge used a sample of 71,556 questionnaires focussed on "Big Five" personality traits (i.e. *openness to experience*, *conscientiousness*, *extraversion*, *agreeableness*, and *neuroticism*) and investigated the relationship between users' personality scores and their Facebook *status updates* – short texts regarding a user's thoughts, feelings, activities – using natural language processing methods. A key result of this work is that personality measures derived from applying natural language processing methods to status updates agreed (i.e. moderate to strong positive correlation) with self-reported personality. [PubMed: 25365036]

38**. Jashinsky J, Burton SH, Hanson CL, West J, Giraud-Carrier C, Barnes MD, Argyle T. Tracking suicide risk factors through Twitter in the US. Crisis. 2014; 35(1):51–9. Noting a lack of effective surveillance infrastructure for suicide, researchers at Brigham Young University developed a method for identifying and then geocoding tweets that could be indicative of suicide risk factors. First, a set of terms related to suicide risk factors (e.g. *impulsivity*, *suicidal ideation*, *self harm*) where iteratively identified and tested. Second, this list of terms (in addition to simple rules) was used to query the Twitter Application Programming Interface, with data collected for a three month period in 2012. The suicide risk factor tweets identified over this period were then geolocated, and correlated with United States vital statistics, yielding a strong positive correlation between suicide risk-factor related language on Twitter and national suicide statistics. [PubMed: 24121153]

39**. De Choudhury, M.; Counts, S.; Horvitz, E. Social media as a measurement tool of depression in populations. In: Davis, HC.; Halpin, H.; Pentland, A.; Bernstein, M.; Adamic, LA., editors. Web Science 2013 (co-located with ECRC), WebSci '13; Paris, France. May 2–4, 2013; ACM; 2013. p. 47-56.De Choudhury et al. developed a crowdsourcing solution to the problem of developing a ground truth data set. First, annotators were recruited from Amazon Mechanical Turk and required to take a CES-D (Center for Epidemiologic Studies Depression Scale) test then asked a series of question regarding their history of depression and current depression status. Second, the Mechanical Turkers who completed the questionnaire were asked for their Twitter user name, which was then used (with the consent of the Turkers) to pull their Twitter feed, resulting in a ground truth depressed/not-depressed data set. Third, a machine learning classifier was trained on the depressed/not-depressed data using features derived from both the tweet text and network features (e.g. number of followers). Fourth, the classifier was applied to a large data set of geolocated Twitter data from the United States, yielding a strong positive correlation with Centers for Disease Control depression statistics. URL http://doi.acm.org/10.1145/2464464.2464480

40. De Choudhury, M.; Counts, S.; Horvitz, E. Predicting postpartum changes in emotion and behavior via social media. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '13; New York, NY, USA: ACM; 2013. p. 3267-3276.http://doi.acm.org/10.1145/2470654.2466447

41. De Choudhury, M.; Gamon, M.; Counts, S.; Horvitz, E. Predicting depression via social media. In: Kiciman, E.; Ellison, NB.; Hogan, B.; Resnick, P.; Soboroff, I., editors. Proceedings of the Seventh International Conference on Weblogs and Social Media, ICWSM 2013; Cambridge, Massachusetts, USA. July 8–11, 2013; The AAAI Press; 2013. http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6124

42. De Choudhury, M.; Counts, S.; Horvitz, E.; Hoff, A. Characterizing and predicting postpartum depression from shared facebook data. In: Fussell, SR.; Lutters, WG.; Morris, MR.; Reddy, M., editors. Computer Supported Cooperative Work, CSCW '14; Baltimore, MD, USA. February 15–19, 2014; ACM; 2014. p. 626-638.http://doi.acm.org/10.1145/2531602.2531675

43. De Choudhury, M.; De, S. Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In: Adar, E.; Resnick, P.; Choudhury, MD.; Hogan, B.; Oh, A., editors. Proceedings of the Eighth International Conference on Weblogs and Social Media, ICWSM 2014; Ann Arbor, Michigan, USA. June 1–4, 2014; The AAAI Press; 2014. http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8075

44. Balani, S.; De Choudhury, M. Detecting and characterizing mental health related self-disclosure in social media. In: Begole, B.; Kim, J.; Inkpen, K.; Woo, W., editors. Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, Seoul, CHI 2015 Extended Abstracts; Republic of Korea. April 18 – 23, 2015; ACM; 2015. p. 1373-1378.http://doi.acm.org/10.1145/2702613.2732733

45. Tanana, M.; Hallgren, K.; Imel, Z.; Atkins, D.; Smyth, P.; Srikumar, V. Recursive neural networks for coding therapist and patient behavior in motivational interviewing. Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; Denver, Colorado: Association for Computational Linguistics; 2015. p. 71-79.http://www.aclweb.org/anthology/W15-1209

46. Rouhizadeh, M.; Sproat, R.; van Santen, J. Similarity measures for quantifying restrictive and repetitive behavior in conversations of autistic children. Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; Denver, Colorado: Association for Computational Linguistics; 2015. p. 117-123.http://www.aclweb.org/anthology/W15-1214

47**. Coppersmith, G.; Dredze, M.; Harman, C. Quantifying mental health signals in Twitter. Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; Baltimore, Maryland, USA: Association for Computational Linguistics; 2014. p. 51-60.Coppersmith et al. applied natural language processing techniques to English language Twitter data with the goal of deriving insights into four mental health conditions (post-traumatic stress disorder, depression, bipolar disorder, and seasonal affective disorder). The chief contribution of this paper is the introduction of a novel method for the identification of tweets likely to be associated with a particular psychiatric diagnosis. Utilising data derived from the Twitter Application Programming Interface, the researchers leveraged pattern matching techniques to identify tweets containing self disclosed diagnoses (e.g. *I was diagnosed with PTSD*). The resulting tweets were then manually analysed in order to remove irrelevant data (e.g. "joke" tweets, obviously sarcastic tweets). Next, the most recent 3,200 tweets from each user identified in the previous step were retrieved using the Twitter Application Programming Interface, resulting in a dataset (for all four mental health conditions) of almost 3 million tweets. URL http://www.aclweb.org/anthology/W/W14/W14-3207

48. Coppersmith, G.; Dredze, M.; Harman, C.; Hollingshead, K. From ADHD to SAD: Analyzing the language of mental health on Twitter through self-reported diagnoses. Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; Denver, Colorado: Association for Computational Linguistics; 2015. p. 1-10.http://www.aclweb.org/anthology/W15-1201

49. Mitchell, M.; Hollingshead, K.; Coppersmith, G. Quantifying the language of schizophrenia in social media. Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; Denver, Colorado: Association for Computational Linguistics; 2015. p. 11-20.http://www.aclweb.org/anthology/W15-1202

50. Preoţiuc-Pietro, D.; Eichstaedt, J.; Park, G.; Sap, M.; Smith, L.; Tobolsky, V.; Schwartz, HA.; Ungar, L. The role of personality, age, and gender in tweeting about mental illness. Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; Denver, Colorado: Association for Computational Linguistics; 2015. p. 21-30.http://www.aclweb.org/anthology/W15-1203

51. Resnik, P.; Armstrong, W.; Claudino, L.; Nguyen, T.; Nguyen, V-A.; Boyd-Graber, J. Beyond LDA: Exploring supervised topic modeling for depression-related language in Twitter. Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; Denver, Colorado: Association for Computational Linguistics; 2015. p. 99-107.http://www.aclweb.org/anthology/W15-1212

52. Mowery, D.; Bryan, C.; Conway, M. Towards developing an annotation scheme for depressive disorder symptoms: A preliminary study using Twitter data. Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; Denver, Colorado: Association for Computational Linguistics; 2015. p. 89-98.http://www.aclweb.org/anthology/W15-1211

53. Schwartz, HA.; Eichstaedt, J.; Kern, ML.; Park, G.; Sap, M.; Stillwell, D.; Kosinski, M.; Ungar, L. Towards assessing changes in degree of depression through Facebook. Proceedings of the

Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; Baltimore, Maryland, USA: Association for Computational Linguistics; 2014. p. 118-125.http://www.aclweb.org/anthology/W/W14/W14-3214

54. Lopez, S. The Oxford Handbook of Positive Psychology. Oxford University Press; 2011.
55. Kern ML, Eichstaedt JC, Schwartz HA, Dziurzynski L, Ungar LH, Stillwell DJ, Kosinski M, Ramones SM, Seligman MEP. The online social self: an open vocabulary approach to personality. Assessment. 2014; 21(2):158–69. [PubMed: 24322010]
56. Eichstaedt JC, Schwartz HA, Kern ML, Park G, Labarthe DR, Merchant RM, Jha S, Agrawal M, Dziurzynski LA, Sap M, Weeg C, Larson EE, Ungar LH, Seligman MEP. Psychological language on Twitter predicts county-level heart disease mortality. Psychol Sci. 2015; 26(2):159–69. [PubMed: 25605707]
57. McKee R. Ethical issues in using social media for health and health care research. Health Policy. 2013; 110(2–3):298–301. [PubMed: 23477806]
58. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ. 2001; 323(7321):1103–5. [PubMed: 11701577]
59. Wells DM, Lehavot K, Isaac ML. Sounding off on social media: The ethics of patient storytelling in the modern era. Acad Med. 2015; 90(8):1015–9.10.1097/ACM.0000000000000668 [PubMed: 25692559]
60. Lunnay B, Borlagdan J, McNaughton D, Ward P. Ethical use of social media to facilitate qualitative research. Qual Health Res. 2015; 25(1):99–109.10.1177/1049732314549031 [PubMed: 25212856]
61. Conway M. Ethical issues in using Twitter for public health surveillance and research: developing a taxonomy of ethical concepts from the research literature. J Med Internet Res. 2014; 16(12):e290. [PubMed: 25533619]

62**. Vayena E, Salathé M, Madoff LC, Brownstein JS. Ethical challenges of big data in public health. PLoS Comput Biol. 2015; 11(2):e1003904. Vayena et al. focus on ethical issues involved in digital disease detection generally, rather than mental health in particular, and points out that the national and international legislation governing public health surveillance and research do not necessarily address ethical challenges posed by the use of consumer generated data for applications in public health. The researchers identify three clusters of ethical challenges that require attention. First, *context sensitivity*, including the need to distinguish between commercial and public health uses of consumer generated data. Second, *nexus of ethics and methodology*, including the need to develop robust methodologies in order to, for example, reduce the possibility of inaccurate reports of disease prevalence. Third *legitimacy requirements*, including the development of best practice community standards for digital disease detection. [PubMed: 25664461]

Author Manuscript
Author Manuscript
Author Manuscript

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