A Case Study in Belief Surveillance, Sentiment Analysis, and Identification of Informational Targets for E-Cigarettes Interventions

Lourdes S. Martinez
Center for Human Dynamics in the Mobile Age
School of Communication
San Diego State University
San Diego, CA USA
lsmartinez@sdsu.edu

Ming-Hsiang Tsou,
Center for Human Dynamics in the Mobile Age
Department of Geography
San Diego State University
San Diego, CA USA
mtsou@sdsu.edu

Brian H. Spitzberg
Center for Human Dynamics in the Mobile Age
School of Communication
San Diego State University
San Diego, CA USA
spitz@sdsu.edu

ABSTRACT

To illuminate understanding of how social media can be leveraged to glean insights into public health issues such as e-cigarette use, we use a social media analytics and research testbed (SMART) dashboard to observe Twitter messages and follow content about e-cigarettes in different cities across the U.S. Our case studies indicate that the majority of e-cigarette tweets are positive (68%), which represents a potential problem for public health. Stigma plays the most important roles in both confirmed and rejected messages for e-cigarettes. We also noticed that some advocates of e-cigarettes might be hybrid human-bot accounts (or multiple users using one account). Our key findings demonstrate the use of the SMART dashboard as a means of public health-related belief surveillance, and identification of campaign targets and informational needs of different communities in real-time. Future uses of this tool include monitoring social messages about e-cigarettes for combating the spread of tobacco-related misinformation and disinformation, and detecting and targeting informational needs of communities for intervention.

CCS CONCEPTS

• **Human-centered computing → Visualization**; *Visual application domains*; Visual analytics • **Social and professional topics → User characteristics**; *Geographic characteristics* • **Social and professional topics → Computing/technology policy**; *Surveillance*; Government surveillance

KEYWORDS

Social Media Analytics, Twitter, E-cigarettes, Geo-target, Spatiotemporal

ACM Reference format:
Lourdes S. Martinez, Ming Hsiang-Tsou, Brian H. Spitzberg. 2019. A Case Study in Belief Surveillance, Sentiment Analysis, and Identification of Informational Targets for E-cigarettes Interventions. In *International Conference on Social Media & Society (SMSociety[1] '19), July 19-21, 2019, Toronto, ON, Canada.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3328529.3328540

1 INTRODUCTION

Social media technology has ushered an era enabling unprecedented opportunities for individuals to connect with one another. This interconnection, however, has arrived with exigencies related to the misuse and abuse of user data gathered through social media platforms [1, 2], and the emergence of social media platforms as avenues for spreading misinformation (e.g., inaccurate information) and disinformation (e.g., deception information) [3-6]. This is particularly problematic when such content becomes memetic and virally spreads across social networks [7-9]. Bots, in particular, present a challenge to reversing trends of the sense of privacy and public trust in social media content. From a privacy perspective, bots are designed to mimic authentic human activity, and may use online user data to be

SMSociety '19, July 19–21, 2019, Toronto, ON, Canada

ACM ISBN 978-1-4503-6651-9/19/07. . . $15.00
https://doi.org/10.1145/3328529.3328540

effectively engaging [10-12]. There are various reasons why users may not wish their data to be used to guide activity of bots in this manner. From a trust perceptive, while not all bots are designed with veiled intent or for nefarious purposes [13], public health experts and scholars are expressing increasing concerns regarding the potential for bots to disseminate misinformation and disinformation related to public health issues [10]. Trust in health information suffers when misinformation and disinformation proliferate to a point where significant correction is required [1], and possibly fuels perceptions that online health information is neither reliable nor accurate [14, 15].

Recent theorizing in the communication discipline suggests that in order to understand when and how public health-related online information (including misinformation and disinformation) is likely to virally spread online across a social network, researchers need to consider human processes and how they employ unique properties of social media in realspace [7, 16]. The multilevel model of meme diffusion (M^3D) is distinctively positioned for this purpose.

Answering the question of what drives the diffusion of information is an ongoing objective of various fields and theories. Distinct models emphasize different variables: aspects of the message or meme itself; the influential sources from which such messages derive; the structural features or human dynamics of the social networks to which such messages are sent; the societal and cultural dynamics that contextualize such messages; and the geotechnical context surrounding such dynamics. Given that research identifies features at each of these levels of analysis, it follows that a multi-level approach is likely needed to fully model such phenomena. The multilevel model of meme diffusion (M^3D) aims to provide such a framework. The M^3D proposes that sets of variables influence the diffusion of a message or meme. Following Dawkins' [17] proposal that memes are comparable to genes in that they transfer information from one person to another, Spitzberg [18-20] further conjectured several conditions in which memes are likely to spread. Memes that are more novel and affect-laden travel farther and faster. Their spread is further facilitated when they are shared by users who are considered more credible, trustworthy, and competent. Network characteristics can also help or hinder spread of memes. For example, social networks that are internally homophilous but externally diverse (e.g., many bridgers) help spread memes faster and further. Additionally, the lack of counterarguments in the larger information environment as well as access to technology among a range of proximal users also promotes the diffusion of memes.

Our web-based social media analytics research testbed (SMART) dashboard includes geotargeted social media (specifically Twitter) application programming interfaces (APIs) that allows for real-time tracking of various topics (URL: http://vision.sdsu.edu/hdma/smart), and has been previously used to monitor a range of topics using keywords to gather data on disease outbreaks, drug abuse, and emergencies related to natural disasters [21]. The SMART dashboard is also well-suited to investigate many of the variables hypothesized by the M^3D. The purpose of the current article is to demonstrate the utility of the SMART dashboard for examining and tracking social media messages in a previously unexplored context: e-cigarettes.

2 BACKGROUND

E-cigarettes (or "electronic cigarettes") represent a type of electronic nicotine delivery systems (ENDS), which have experienced increasing use by adolescents [22] and emerging adults [23, 24]. E-cigarette use is a growing public health concern because these devices may contain potentially toxic chemicals [25] and their long-term effects on health remain unestablished [26]. For young people, e-cigarettes may act as a gateway into use of combustible tobacco products [27]. Young people are also heavily engaged users of social networking sites such as Twitter [28], which are often viewed as important information sources [29, 30] and settings for socialization [31]. In addition, a lack of distinction between online and offline life among young individuals is likely to increase impact of messages shared over social media [32]. However, more research is needed to examine the types of messages about e-cigarettes that are shared over social media, who is sharing these messages (e.g., bots versus authentic humans), and the role of e-cigarettes advocates [33] and their promotion of these devices. We introduce two case studies using the SMART dashboard to examine content about e-cigarettes and sources of these messages on Twitter, as well as how they operate in real-time.

3 CASE STUDIES IN E-CIGARETTES

The SMART dashboard is constructed with several data mining programs, GIS methods, and geo-targeted social media APIs in order to monitor topics of interest through space and across time. Capabilities of this tool permit researchers to generate visualizations, and perform descriptive and predictive analyses of these topics in various

U.S. cities across time. Stakeholders, such as government officials and those involved in healthcare delivery and first response, can easily access this tool. We note the following features of the SMART dashboard to further emphasize its unique and important capabilities:

1. Collect and update social media messages daily, along with their spatial attributes and geographic patterns of diffusion in different cities.
2. Present evolution of social media trends (daily, weekly, monthly) over time as they occur in real-time.
3. Filter data to extract noise and errors to improve accuracy of analyses and tracking of social media messages.
4. Display temporal trends of social media messages by individual cities or by aggregation of data across all listed cities.
5. Provide insight into social media messages and how they differ between cities, which may be used by local health agencies and organizations to map geographical hot spots or areas in need of intervention.

Prior research has already demonstrated the utility of the SMART dashboard for collecting and analyzing social media messages in other contexts (e.g., Flu, Whooping Cough, Wildfire, Drugs, and Aztecs) [18, 34-47], and details of the SMART dashboard's original development are available elsewhere [48, 49]. In the present study, we expand on this prior work by examining the use of the SMART dashboard to examine social media messages in the context of e-cigarettes. By entering the following list of keywords [50, 51] into the SMART dashboard, we were able to collect a total of 193,051 tweets between October 2015 and February 2016 across all regions in the U.S.: Vaping, Vape, Vaper, Vapers, Vapin, Vaped, Evape, Vaporing, e-cig*, ecig*, e-pen, epen, e-juice, ejuice, e-liquid, or eliquid. The SMART dashboard retained only tweets that included at least one of these listed keywords. The filtering and data cleaning functions of the SMART dashboard are based on past work [21], and are summarized in Fig. 1 and adapted for the current study.

Figs. 2-3 further detail these and additional functions of the SMART dashboard. The following sections introduce two case studies in the context of e-cigarettes to illustrate how the SMART dashboard can be used to study public health topics. These examples focus on public perceptions and beliefs of e-cigarettes guided by concepts of infodemiology and infoveillance [52]. Infodemiology refers to scientific approaches in the study of online content (collected and analyzed in real-time) used to draw insights in order to advise public health and public policy. Infoveillance uses infodemiological methods with the objective of surveillance.

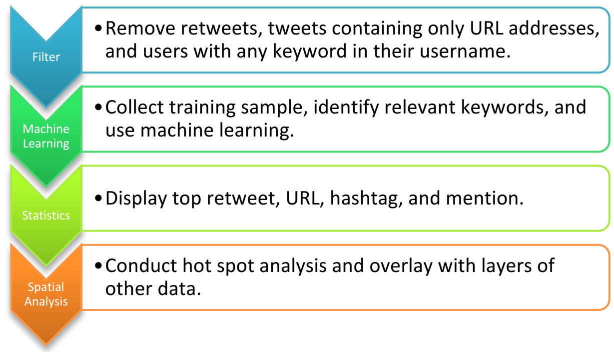

Figure 1: Data Filter and Cleaning Procedures in SMART Dashboard (adapted [21]).

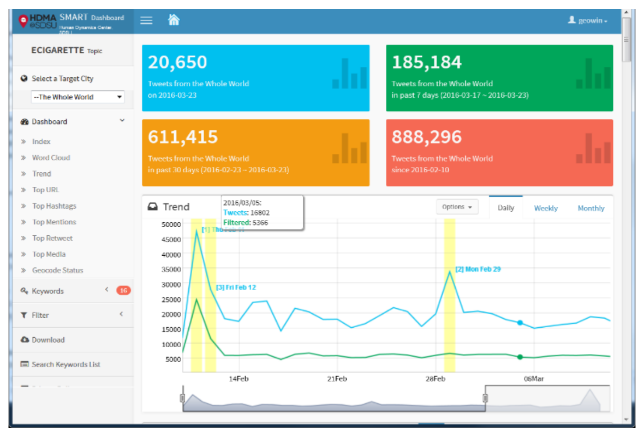

Figure 2: The Screen Shot of SMART Dashboard for E-cigarettes.

Although surveillance represents an important public health activity, traditional approaches of disease detection are typically expensive and can be labor intensive. In contrast, social media content offers data that can be collected and analyzed in real-time, and may provide public health practitioners and researchers with an alternative way to monitor and survey disease outbreaks. We elected to discuss these two case studies to emphasize the advantages and opportunities for disease surveillance offered by tools employing social media analytics, including the SMART dashboard.

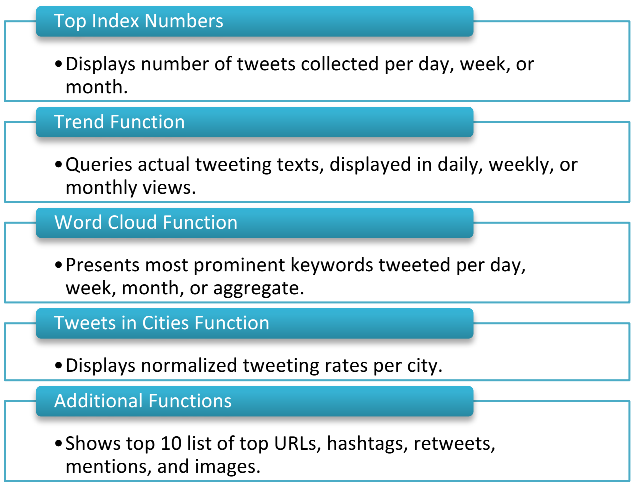

Figure 3: Key Features for Interactive Query and Visualizations in SMART dashboard [21].

3.1 Belief Surveillance

Figs. 4-5 show results of a sentiment analysis for a random sample of tweets (N=973) collected between October 2015 and February 2016 across the U.S. We performed a general sentiment analysis (Fig. 4) and an analysis specifically examining whether tweets confirmed or rejected commonly held beliefs about e-cigarettes (e.g., stigma associated with e-cigarettes, perceived harmfulness, capacity for generating second-hand smoke, helpfulness as a cessation tool, versatility, and potential for addiction) (Fig. 5). The general sentiment analysis examined whether the tweet conveyed content that struck a tone that was positive (approval of e-cigarettes), negative (disapproval of e-cigarettes), neutral (neither approving or disapproving e-cigarettes), ambiguous (both approving and disapproving of e-cigarettes), or other (nonsensical or incomprehensible) [40]. The categories for a more specific analysis of sentiment included: a) the effectiveness or ineffectiveness of e-cigarette use as a cessation tool; (b) increased or reduced addictiveness of e-cigarettes compared to traditional cigarettes; (c) presence or lack of stigma regarding e-cigarettes; (d) whether e-cigarettes cause or reduce 2nd-hand smoke; (e) if e-cigarettes produce beneficial or harmful effects on health; (f) freedom or restrictions on users and when or where they can vape; and (g) the general satisfaction or dissatisfaction from using e-cigarettes instead of traditional e-cigarettes. We can use a graph similar to this to compare sentiment related to e-cigarettes from official regional health census data, as well as data from Monitoring the Future, a National Institute on Drug Abuse and NIH ongoing study of young adult attitudes, values and behaviors [53], capturing perceptions of e-cigarettes as a cessation tool, along with personal disapproval of and perceived risk from regularly using e-cigarettes.

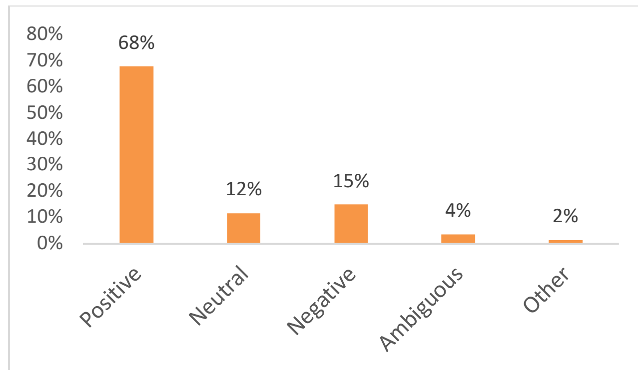

Figure 4: Sentiment Analysis (N=973) [40]

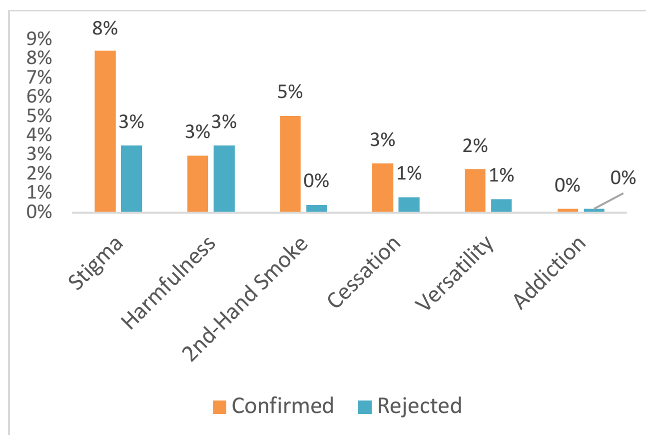

Figure 5: Sentiment Analysis: Confirmation and Rejection of Common Beliefs About E-Cigarettes (N=973) [40]

Fig. 4 indicated that the majority of e-cigarettes tweets are positive (68%), indicating support for the use of e-cigarettes. Such pro-vaping content could have a very negative impact for public health. Fig. 5 illustrated that "Stigma" and "Second-Hand Smoke" are the major reasons in the sampled tweets that support the use of e-cigarettes. On the other hand, "Harmfulness" and "Stigma" are the primarily expressed reasons to reject e-cigarettes. "Stigma" plays the most important roles in both confirmed and rejected messages for e-cigarettes. Our interactive SMART dashboard can provide practitioners with a way to collect social media messages in real-time to monitor and visualize trends in e-cigarettes sentiment. The daily, weekly, and monthly monitoring functions of the SMART dashboard offer public health authorities at state and local levels a tool for collecting data on beliefs about e-cigarettes. Sentiment analysis of these data can indicate how beliefs are changing over time, point toward opportunities for intervention, and present an additional way to evaluate existing intervention efforts. Using the M^{3}D's meme-level concepts to distinguish which types of tweets are likely to spread, we can also examine the content of tweets for these characteristics before they become viral. For example, the finding that stigma was central to both confirmatory and disconfirmatory messages related to e-cigarettes suggests it may help fuel a larger debate, which is likely to garner attention among users and increase the likelihood of spreading across social networks. In this way, the identification of tweets before they potentially become viral can help thwart the spread of health misinformation, or even suggest message designs that could effectively counter-argue such viral trends. This would also allow public health practitioners to get ahead of the conversation online with argumentation that can potentially slow or stop tweets that endorse e-cigarettes and that exhibit memetic potential from becoming viral.

3.2 Promotion and Advocates

A second case study uses the SMART dashboard to identify proponents of behaviors that may undermine public health goals. In the context of e-cigarettes, advocates [33] on social media may voice positions promoting the use of e-cigarettes that could shape views and risk perceptions of vulnerable populations, including younger populations. Fig. 6 shows a summary of activity rates for authentic human accounts acting as advocates of e-cigarettes generated by the SMART dashboard, including network size and daily, weekly, and monthly activity. The daily rates are calculated using the whole lifespan of user accounts. The weekly rates are based on the last seven days (most recently). The monthly rates are based on the last 30 days. We found that most advocates' daily post numbers are between 3 and 53. Their daily, last seven days, and last 30 days rates are very consistent except the #4 account, which might be a hybrid human-bot account (or multiple users using one account) with 355 average daily posts. Also, the #4 account was created within 30 days of our collection period (missing the last 30 days activity rates). The SMART dashboard also provides data for word clouds, which can offer insight into the most prominently featured words used by advocates. Figs. 7-8 illustrate two sample profiles for advocates, including the five most recent tweets posted by the user, and a word cloud generated using the last 3,200 tweets shared. Fig. 7 illustrates the activities from a human advocate (#1) and Fig. 8 illustrates the activities from a potential cyborg account (#4).

1	#	Twitter Account Start Date	Followers	Daily Activity Rate	Last 7 Days Activity Rate	Last 30 Days Activity Rate
2	5	9/23/2015	**8639**	**10**	5	8
3	2	4/24/2015	**6024**	**8**	10	17
4	6	4/10/2016	**2435**	**40**	25	30
5	4	6/22/2016	**2286**	**355**	358	**na**
6	1	4/27/2016	**1078**	**53**	58	39
7	10	11/16/2015	**887**	**13**	17	15
8	14	4/17/2016	**666**	**37**	44	31
9	8	1/19/2014	595	6	10	7
10	13	2/25/2016	520	26	20	25
11	7	10/24/2015	508	12	8	10
12	3	11/19/2015	492	17	16	19
13	9	4/17/2012	348	10	2	3
14	11	10/2/2014	278	3	5	3
15	12	8/23/2012	143	3	1	3

Figure 6: Activity Rate for Advocates Twitter Accounts (ranked by the numbers of followers).

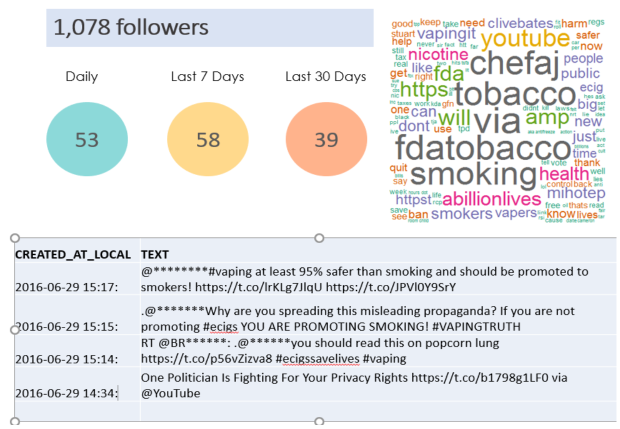

Figure 7: Activity Rate for a Normal Advocates Twitter Account (#1).

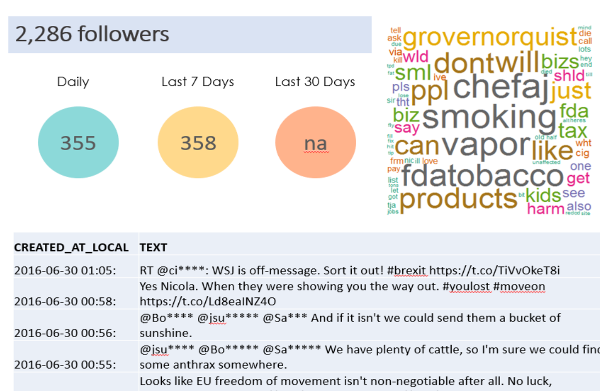

Figure 8: Activity Rate for a Cyborg Advocates Twitter Account (#4).

4 CONCLUSION

Social media analytics provide opportunities for collecting data that can be used in disease surveillance and improving understanding of behavioral determinants of diseases. The SMART dashboard provides one tool to advance some of these opportunities by serving as a means of public health-related belief surveillance, bot detection, and identification of campaign targets and informational needs of different communities in real-time. This project extends research in the area of e-cigarettes and social media analytics by gathering geo-tagged tweets, and offering spatiotemporal insight into social media messages posted about e-cigarettes. Specifically, we are able to track beliefs and risk perceptions about e-cigarettes, bot activity, and campaign targets with a spatiotemporal view. Linking time and place together using the M^3D model in this manner allows discovery of important patterns that illuminate understanding of disease transmission and social media activities. We have demonstrated the use of the SMART dashboard in this capacity within the context of e-cigarettes as an example for how public health practitioners and researchers may consider using this tool in the future for other public health issues. The individual level of M^3D model can be used to study the motivation and skills of the e-cigarette advocates.

The use of the SMART dashboard for these purposes, however, does present certain challenges that merit discussion. The first challenge relates to the issue of privacy and its protection, which remains a significant concern for all social media analytic tools. Throughout the process of developing the SMART dashboard, our team has strived to protect the privacy of individuals using social media. Specifically, we only gather public tweets made available through the public Twitter APIs. We also provide a Privacy Policy inviting concerned users to reach out: "*If you have any concerns about the privacy issues in our web applications, please Email us. After verify your information, we will remove specific social media contents based on your requests.*" One option for enhancing privacy protection is to assign anonymous IDs to all users. This approach, however, could undermine research seeking to understand social networks and the attributes that contribute to the creation and diffusion of social media messages. Such challenges will require future researchers to weigh the suitability of social media messages as a data source with the need to protect user privacy.

5 ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1416509, project titled "Spatiotemporal Modeling of Human Dynamics Across Social Media and Social Networks" and No. 1634641, "Integrated Stage-based Evacuation with Social Perception Analysis and Dynamic Population Estimation". Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors are grateful for contributions from other HDMA team members in the development of SMART dashboard.

6 REFERENCES

[1] Jon-Patrick Allem, and Emilio Ferrara. 2016. The importance of debiasing social media data to better understand e-cigarette-related attitudes and behaviors. *Journal of Medical Internet Research* 18, 8 (2016), e219. doi:10.2196/jmir.6185

[2] Marc P. Armstrong, Ming-Hsiang Tsou, and Dara E. Seidl. 2018. Geoprivacy. In *Comprehensive Geographic Information Systems*. Amsterdam, Elsevier, 415–30.

[3] Alessandro Bessi, Mauro Coletto, George A. Davidescu, Antonia Scala, Guido Caldarelli, and Walter Quattrociocchi. 2015. Science vs conspiracy: Collective narratives in the age of misinformation. *PLoS One* 10, 2 (2015), e0118093. doi:10.1371/journal.pone.0118093

[4] Wen-Ying Sylvia Chou., April Oh, and William M.P. Klein. 2018. Addressing health-related misinformation on social media. *JAMA: Journal of the American Medical Association* 320, 23 (2018), 2417–2418. doi:10.1001/jama.2018.16865

[5] Jennifer Kavanagh, and Michael D. Rich. 2018. *Truth decay: An initial exploration of the diminishing role of facts and analysis in American public life.* Santa Monica, CA: RAND Corp. https://www.rand.org/content/dam/rand/pubs/research_reports/RR2300/RR2314/RAND_RR2314.pdf

[6] Lee Rainie, Janna Anderson, and Jonathan Albright. 2017. The future of free speech, trolls, anonymity and fake news online. Washington DC: Pew Research Center. http://www.pewinternet.org/2017/03/29/the-

future-of-free-speech-trolls-anonymity-and-fake-news-online/

[7] Jieun Shin, Lian Jian, Kevin Driscoll, and Francois Bar. 2018. The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior* 83 (2018), 278–287. doi: 10.1016/j.chb.2018.02.008

[8] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. 2016. The spreading of misinformation online. In *Proceedings of the National Academy of Sciences of the United States of America* 113, 3 (2016), 554-559. doi:10.1073/pnas.1517441113

[9] Bernie Garrett, Sue Murphy, Shahin Jamal, Maura MacPhee, Jillian Reardon, Winson Cheung, Emilie Mallia, Cathryn Jackson. 2019. Internet health scams—Developing a taxonomy and risk-of-deception assessment tool. *Health & Social Care in the Community* 27*,* 1 (2019), 226–240. doi: 10.1111/hsc.12643

[10] David A. Broniatowski, Amelia M. Jamison, SiHua Qi, Lulwah AlKulaib, Tao Chen, Adrian Benton, Sandra C. Quinn, and Mark Dredze. 2018. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *American Journal of Public Health* 108*,* 10 (2018), 1378–1384. doi: 10.2105/AJPH.2018.304567

[11] Massimo Stella, Emilio Ferrara, and Manlio De Domenico. 2018. Bots increase exposure to negative and inflammatory content in online social systems. In *Proceedings of the National Academy of Sciences of the United States of America* 115*,* 49 (2018), 12435–12440. doi: 10.1073/pnas.1803470115

[12] Jeannette Sutton. 2018. Health communication trolls and bots versus public health agencies' trusted voices. *American Journal of Public Health* 108*,* 10 (2018), 1281–1282. doi:10.2105/AJPH.2018.304661

[13] Carolina Alves de Lima Salge, and Nicholas Berente. 2017. Computing ethics is that social bot behaving unethically? A procedure for reflection and discourse on the behavior of bots in the context of law, deception, and societal norms. *Communications of the ACM* 60*,* 9 (2017), 29–31. doi:10.1145/3126492

[14] Jon-Patrick Allem, Emilio Ferrara, Sree Priyanka Uppu, Tess Boley, and Jennifer B. Unger. 2017. E-cigarette surveillance with social media data: Social bots, emerging topics, and trends. *Journal of Medical Internet Research* 19*,* 12 (2017), 1. doi:10.2196/publichealth.8641

[15] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. 2016. The rise of social bots. *Communications of the ACM* 59, 7 (2016), 96–104. doi:10.1145/2818717

[16] Carlo Kopp, Kevin B. Korb, and Bruce I. Mills. 2018. Information-theoretic models of deception: Modelling cooperation and diffusion in populations exposed to "fake news." *PLoS One* 13*,* 11 (2018), 1–35. doi:10.1371/journal.pone.0207383

[17] Richard Dawkins. 1976. *The selfish gene*. New York, NY: Oxford University Press.

[18] Adiyana Sharag-Eldin, Xinyue Ye, and Brian Spitzberg. 2018. Multilevel model of meme diffusion of fracking through Twitter. *Chinese Sociological Dialogue* 3*,* 1 (2018), 17-43. doi: 10.1177/2397200917752646

[19] Brian H. Spitzberg. 2014. Toward a model of meme diffusion (M^3D). *Communication Theory* 24*,* 3 (2014), 311-339. doi:10.1111/comt.12042

[20] Michael P. Schlaile, Theresa Knausberg, Matthias Mueller, Johannes Zeman. 2018. Viral ice buckets: A memetic perspective on the ALS ice bucket challenge's diffusion. *Cognitive Systems Research* 52, (2018), 947-969. doi:10.1016/j.cogsys.2018.09.012

[21] Ming-Hiang Tsou, Chin-Te Jung, Chris Allen, Jiue-An Yang, Jean-Mark Gawron, Brian H. Spitzberg, and Su Han. 2015. Social media analytics and research test-bed (SMART dashboard). In Proceedings of *the 2015 International Conference on Social Media & Society,* (2015), p. 2). ACM. doi: 10.1145/2789187.2789196

[22] Rene Arrazola, Tushar Singh, Catherine G. Corey, Corinne G. Husten, Linda J. Neff, Benjamin J. Apelberg, Rebecca E. Bunnell, Conrad J. Choiniere, Brian A. King, Shanna Cox, Tim McAfee, and Ralph. K. Caraballo. 2015. Tobacco use among middle and high school students-United States, 2011-2014. *MMWR. Morbidity and Mortality Weekly Report* 64, 14 (2015), 381-385.

[23] Kelvin Choi, and Jean Forster. 2013. Characteristics associated with awareness, perceptions, and use of electronic nicotine delivery systems among young US Midwestern adults. *American Journal of Public Health* 103*,* 3 (2013), 556-561. doi: 10.2105/AJPH.2012.300947

[24] Brian A. King, Suhana Alam, Gabbi Promoff, Rene Arrazola, and Shanta Dube. 2013. Awareness and ever-use of electronic cigarettes among US adults, 2010–2011. *Nicotine &*

Tobacco Research 15, 9 (2013), 1623-1627. doi: 10.1093/ntr/ntt013
[25] Nathan K. Cobb, M. Justin Byron, David B. Abrams, Peter G. Shields. 2010. Novel nicotine delivery systems and public health: the rise of the "e-cigarette". *American Journal of Public Health* 100, 12 (2010), 2340-2342. doi: 10.2105/AJPH.2010.199281
[26] Rachel Grana, Neal Benowitz, and Stanton A. Glantz. 2014. E-cigarettes a scientific review. *Circulation* 129, 19 (2014), 1972-1986. doi: 10.1161/CIRCULATIONAHA.114.007667
[27] Jessica L. Barrington-Trimis, Robert Urman, Adam M. Leventhal, W. James Gauderman, Tess Boley Cruz, Tamika D. Gilreath, Steve Howland Jennifer B. Unger, Kiros Berhane, Jonathan M. Samet, and Robert McConnell. 2016. E-cigarettes, cigarettes, and the prevalence of adolescent tobacco use. *Pediatrics*, (2016), e20153983. doi: 10.1542/peds.2015-3983
[28] Maeve Duggan, and Joanna Brenner. 2013. *The demographics of social media users, 2012* (Vol. 14). Washington, DC: Pew Research Center's Internet & American Life Project.
[29] Kyung-Sun Kim, Sei-Ching Sin, and Tien-I Tsai. 2014. Individual differences in social media use for information seeking. *The Journal of Academic Librarianship* 40, 2 (2014), 171-178. doi: 10.1016/j.acalib.2014.03.001
[30] David Westerman, Patric R. Spence, and Brandon Van Der Heide. 2014. Social media as information source: Recency of updates and credibility of information. *Journal of Computer-Mediated Communication* 19, 2 (2014), 171-183. doi: 10.1111/jcc4.12041
[31] Louise Barkhuus, and Juliana Tashiro. 2010. Student socialization in the age of Facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* (2010), p. 133-142. ACM. doi: 10.1145/1753326.1753347
[32] Michelle F. Wright, and Yan Li. 2011. The associations between young adults' face-to-face prosocial behaviors and their online prosocial behaviors. *Computers in Human Behavior* 27, 5 (2011), 1959-1962. doi: 10.1016/j.chb.2011.04.019
[33] Annice Kim, Thomas Miano, Robert Chew, Matthew Eggers, James Nonnemaker. 2017. Classification of Twitter users who tweet about e-cigarettes. *JMIR Public Health and Surveillance* 3, 3 (2017), e63. doi: doi:10.2196/publichealth.8060
[34] Xuan Shi, Bowei Xue, Ming-Hsiang Tsou, Xinyue Ye, Brian H. Spitzberg, Jean-Mark Gawron, Heather Corliss, Jay Lee, and Ruoming Jin. 2018. Detecting events from the social media through exemplar-enhanced supervised learning. *International Journal of Digital Earth, (*2018), 1-15. doi: 10.1080/17538947.2018.1502369
[35] Chris Allen, Ming-Hsiang Tsou, Anoshe Aslam, Anna Nagel, and Jean-Mark Gawron. 2016. Applying GIS and machine learning methods to Twitter data for multiscale surveillance of influenza. *PLoS One* 11, 7 (2016), e0157734. doi:10.1371/journal.pone.0157734
[36] Anoshe A. Aslam, Ming-Hsiang Tsou, Brian H. Spitzberg, Li An, Jean-Mark Gawron, Dipak K. Gupta, K. Michael Peddecord, Anna C. Nagel, Christopher Allen, Jiue-An Yang, & Suzanne Lindsay. 2014. The reliability of tweets as a supplementary method of seasonal influenza surveillance. *Journal of Medical Internet Research* 16, 10 (2014), e250, 1-12. doi: 10.2196/jmir.3532
[37] Su-Yeon Han, Ming-Hsiang Tsou, and Keith C. Clarke. 2018. Revisiting the death of geography in the era of big data: The friction of distance in cyberspace and real space. *International Journal of Digital Earth* 11, 5 (2018), 451-469. doi: 10.1080/17538947.2017.1330366.
[38] Elias Issa, Ming-Hsiang Tsou, Atsushi Nara, Brian H. Spitzberg. 2017. Spatio-temporal characteristics of Twitter data with geotagged and non-geotagged content: Two case studies with topic of flu and *Ted* (movie). *Annals of GIS* 23, 3 (2017), 1-17. doi: 10.1080/19475683.2017.1343257
[39] Ick-Hoi Kim, Chen-Chieh Feng, Yi-Cheng Wang, Brian H. Spitzberg, and Ming-Hsiang Tsou. 2017. Exploratory spatiotemporal analysis in risk communication during the MERS outbreak in South Korea. *The Professional Geographer* 69, 4 (2017), 629-643, doi: 10.1080/00330124.2017.1288577
[40] Lourdes S. Martinez, Sharon Hughes, Eric Walsh-Buhi, and Ming-Hsiang. 2018. "Okay, we get it. You vape": An analysis of geocoded content, context, and sentiment regarding E-cigarettes on Twitter. *Journal of Health Communication* 23, 6 (2018), 550-562: doi: 10.1080/10810730.2018.1493057
[41] Lourdes S. Martinez, Brian H. Spitzberg, Ming-Hsiang Tsou, Elias Issa, E., K. Michael Peddecord. 2017. *Vax Populi*: The social [media] (de)construction of public health policy. Paper presented at the International Communication Association Conference, San Diego, CA.
[42] Anna C. Nagel, Ming-Hsiang Tsou, Brian H. Spitzberg, Li An, Jean-Mark Gawron, Dipak K. Gupta, Jiue-An Yang, Su Han, K. Michael

Peddecord, Suzanna Lindsay, and Mark H. Sawyer. 2013. The complex relationship of real-space events and messages in cyberspace: A case study of influenza and pertussis using tweets. *Journal of Internet Medical Research* 15, 10 (2013), e237, 1-13. doi: 10.2196/jmir.2705

[43] Mario Sanguinet, Brian H. Spitzberg, and Ming-Hsiang Tsou. 2017. *Hashtags, tweets and movie receipts: Social media analytics in predicting box office hits*. Paper presented at the National Communication Association Conference, Dallas, TX.

[44] Chris Allen, Ming-Hsiang Tsou, Anoshe Aslam, Anna Nagel, and Jean-Mark Gawron. 2016. Applying GIS and machine learning methods to Twitter data for multiscale surveillance of influenza. *PloS One*, 11, 7 (2016), 1-10. doi:10.1371/journal.pone.0157734

[45] Zheye Wang, Xinyue Ye. 2018. Social media analytics for natural disaster management. *International Journal of Geographical Information Science* 32, 1 (2018), 49-72. doi: 10.1080/13658816.2017.1367003

[46] Zheye Wang, and Xinyue Ye. 2018. Space, time, and situational awareness in natural hazards: A case study of Hurricane Sandy with social media data. *Cartography and Geographic Information Science*, (2018). doi:10.1080/15230406.2018.1483740

[47] Xinyue Ye, Adiyana Sharag-Eldin, and Brian H. Spitzberg. 2018. Analyzing public opinions on death penalty abolishment. *Chinese Sociological Dialogue,* (2018). doi: 10.1177/2397200918761665

[48] Ming-Hsiang Tsou, Cin-Te Jung, Christopher Allen, Jiue-An Yang, Su Yeon Han, Brian H. Spitzberg, and Jessica Dozier. 2017. Building a real-time geo-targeted event observation (GEO) viewer for disaster management and situation awareness. In M. P. Peterson (Ed.), *Advances in cartography and GIScience: Selections from the International Cartographic Conference 2017* (Series: Publications of the International Cartographic Association, pp. 85-98). New York, NY: Springer.

[49] Jiue-An Yang, Ming-Hsiang Tsou, Chin-Te Jung, Christopher Allen, Brian H. Spitzberg, Jean-Mark Gawron, and Su Yeon Han. 2016. Social Media Analytics and Research Testbed (SMART): Exploring spatiotemporal patterns of human dynamics with geo-targeted social media messages. *Big Data & Society* 3, 1 (2016), 1-19. doi: 10.1177/2053951716652914

[50] Heather Cole-Lewis, Arun Varghese, Amy Sanders, Mary Schwarz, Jillian Pugatch, Erik Auguston. 2015. Assessing electronic cigarette-related tweets for sentiment and content using supervised machine learning. *Journal of Medical Internet Research* 17, 8 (2015), e208. doi:10.2196/jmir.4392

[51] Jidong Huang, Rachel Kornfield, Glen Sczcypka, and Sherry L. Emery. 2014. A cross-sectional examination of marketing of electronic cigarettes on Twitter. *Tobacco Control* 23, 3 (2014). doi:10.1136/tobaccocontrol-2014-051551

[52] Jeong-Nam Kim. 2018. Digital networked information society and public health: problems and promises of networked health communication of lay publics. *Health Communication* 33, 1 (2018), 1–4. doi:10.1080/10410236.2016.1242039

[53] L. Collins, Allison M. Glasser, Haneen Abudayyeh, H., Jennifer L. Pearson, J. and Andrea C. Villanti. 2018. E-cigarette marketing and communication: How e-cigarette companies market e-cigarettes and the public engages with e-cigarette information. *Nicotine & Tobacco Research* 21, 1 (2018), 14-24. doi:10.1093/ntr/ntx284