**Using Twitter to Examine Social Rationales for Vaccine Refusal**

Michael Smith, George Washington University; Mark Dredze, Johns Hopkins University; David Broniatowski, George Washington University

**Abstract**: Vaccines are critical technology to control disease spread, but vaccine refusal compromises efforts towards this end. Recent studies focus on the social rationales underlying vaccine refusal [1], but effective techniques to address these rationales remain unclear. We use big data to analyze these sociotechnical interactions by studying the distribution of the public's rationales regarding vaccine technology in the US. Since Twitter has proven effective in disease surveillance [2] we use it to collect and track social trends about vaccines. We next detail our data collection, then motivate specific illuminatory analyses. Implications for future work are discussed.

To obtain social information about vaccine technology, we employ machine learning (ML) processes to filter Twitter data described in previous work [3]. This yields a) if a tweet is relevant to vaccination or not; b) if a vaccine tweet is of neutral or non-neutral sentiment, and c) if a non-neutral vaccine tweet is of positive or negative sentiment. We also ensure tweets are geotagged using ML [4].

Temporal, spatial, and social analyses may easily be conducted using these data. At a basic level, we have created a map of the US detailing who discusses vaccination and where, at the state and county levels. Next, we will use ML to detect topics automatically. This will show us specific sub-conversations that are pertinent to vaccination within various communities. We will be able to analyze these both spatially and temporally. This will identify when and where new rationales for vaccine refusal appear, and will offer insights into sub-conversations and associated motivations for vaccine refusal. Different ML algorithms also enable automatic identification of social media users' demographics such as gender or age. Combining this information will allow us to pinpoint how vaccine refusal rationales vary across communities, in addition to when and where. Such analyses will contribute towards understanding the interactions between the spread of a disease, the associated social dynamics, and how technology may be used to reduce overall disease prevalence.

References:

[1] J. Downs, W. de Bruin and B. Fischhoff, "Parents' vaccination comprehension and decisions", Vaccine, vol. 26, no. 12, pp. 1595-1607, 2008.

[2] D. Broniatowski, M. Paul and M. Dredze, "National and Local Influenza Surveillance through Twitter: An Analysis of the 2012-2013 Influenza Epidemic", PLoS ONE, vol. 8, no. 12, p. e83672, 2013.

[3] M. Dredze, D. Broniatowski, M. Smith and K. Hilyard, "Understanding Vaccine Refusal: Why We Need Social Media Now", American Journal of Preventive Medicine, 2015.

[4] M. Dredze, M. Paul, S. Bergsma and H. Tran, "Carmen: A Twitter Geolocation System with Applications to Public Health", in AAAI Workshop on Expanding the Boundaries of Health Informatics Using AI (HIAI), 2013.