Social Scientific Frameworks For Guiding the Use of Social Media and Big Data in Hyperlocal Human Dynamics Research

A position paper for Specialist Meeting on Human Dynamics in the Mobile Age, August 11-12 2015

Author: Lourdes S. Martinez, School for Communication, San Diego State University

Several years ago, Chris Anderson made a bold prediction that the end of theory was upon us with the birth of big data. The reasoning behind his argument was that basic correlations will dictate future science and eliminate the need for hypothesis-testing, causal interpretation, and theories in scientific study. However such a statement contains an important non-sequitur: The core of every scientific investigation is the discovery and understanding of causal relationships, and theory helps organize these causal relationships into useful models. Although the advent of social media and big data has ushered in new opportunities for studying human behavior at both the individual level and collective level, contrary to what some have argued is the end of any need for theoretical models, the emergence of social media and big data signals the need for more (not less) use of theory. One area ripe for the exploration of theoretical frameworks in using social media and big data is hyperlocal human dynamics. Human dynamics in a transdisciplinary area of research that seeks to answer questions related to individuals' behavior, interaction, and communication and how patterns that govern these activities change over time. The study of hyperlocal human dynamics examines these questions in the context of specified localized geographical areas. Although a number of innovative frameworks from social science fields have much to offer the study of big data and social media, this paper focuses on two that are of particular relevance to hyperlocal human dynamics: The Social Amplification of Risk Framework (SARF) and the Model of Meme Diffusion (M3D). The first focuses on how risk messages transmitted between source and receiver may become distorted, while the second focuses on the process by which messages (conditional on a variety of features in their environment) can become viral over time.

The SARF seeks to explain the dynamic social processes that create risk amplification or intensification, which happens when hazard events deemed low in risk by experts generate great public concern and strong sociopolitical reactions. One example of socially amplified risk is the recent rise in mistrust over the efficacy of vaccine compliance during childhood harbored by some parent groups in the U.S. Despite comparatively low risk for most children, a rising number of parents are advocating a noncompliance stance with regards to vaccinating their children on schedule, if at all. The SARF also examines the dynamic social processes underlying the reverse outcome. Risk attenuation occurs when the public pays little attention to a hazard event considered by experts to be relatively high risk. Previously observed examples of risk attenuation include comparatively lower public perceptions of risk from significant hazards posed by smoking and climate change. In the context of vaccine noncompliance, parents who choose not to have their children vaccinated have likely experienced risk attenuation regarding the risk of their children contracting infectious diseases. According to the SARF, information about risk events passes through a series of "amplification stations" or points during the transmission process in which a message may be amplified or weakened. As such, amplification stations can alter the amount, emphasis, interpretation, and elaboration of a message, all of which can influence how receivers construe the information. This in part explains the rapid rise of the anti-vaccination movement via social media.

Research on hyperlocal human dynamics may benefit from using the SARF given that the impact of many risk events is likely to be confined to a localized geographical area, and communities within these local areas are likely to have distinguishing characteristics that can shape risk perceptions. As the social processes involved in risk amplification and risk attenuation have the ability to shape both individual and collective behavior, the SARF is particularly attractive for testing new methodologies using social media and big data. For example, in the context of parental vaccine noncompliance, the SARF can be used to guide data collection of individual-level social amplification stations in an effort to identify "risk amplifiers" and "risk attenuators", or social media users who provide commentary on stories or postings that suggest an either higher or lower risk perception regarding vaccine compliance (or noncompliance) than what is reported in the mainstream media. Geotags can be part of this data collection in order to link social media activities of these individuals with geographic locations. Analysis of this data can then be used to determine whether and where risk amplification of vaccine compliance (or risk attenuation of vaccine

noncompliance) is taking place and may serve as early indicators of a need for intervention. The process of data collection can be further facilitated by developing an algorithm designed to quickly identify risk amplifiers and risk attenuators. Visually, social network data also can be aggregated and examined to detect geographical clusters of risk attenuators or risk amplifiers related to vaccines and how they change over time. Lastly, the SARF might be useful in predicting which geographical communities may be at increased risk for infectious disease outbreaks by looking at characteristics of relevant individuals within those communities.

Another framework that may be useful in the area of hyperlocal human dynamics is the Model of Meme Diffusion (M3D). According to M3D, a meme can be any a behavior, symbol, or idea that can be replicated or imitated, and competes with other memes to achieve the widest diffusion through a social system. For example, in early 2015 an anti-vaccination meme comparing forced vaccination to rape went viral over social media, eliciting strong reactions from both sides of the debate. Drawing from earlier theories of social science, the M3D identifies characteristics at various levels that increase a meme's chances for transmission. These levels describe factors that facilitate meme diffusion at the meme level, individual level, social network level, societal level, geo-technical level, and outcome level.

The M3D may be well suited for assessing how far, fast, and for how long a meme travels through social systems. In the context of hyperlocal human dynamics, this model could be used to rapidly diffuse a message through social networks within a community setting. For example, in the event of an infectious disease outbreak in a localized geographical area, public health professionals may develop an emergency message to operate as a meme and design it with characteristics that would accelerate the message's spread and travel speed throughout social networks in at-risk communities. In addition, a visualization of a meme's flow through social media may be a helpful evaluative and predictive tool for hyperlocal purposes. The effectiveness of the message could be evaluated using criteria for assessing meme efficacy including measuring the number of individuals who adopted the message, the number who passed the message along to others, the speed at which the meme spread through a social network, and how long the message persisted in its original form. The spread, speed, and persistency of the meme could be tracked by collecting data on how frequently the meme is posted, shared with others, and receives commentary via social media. This data can also be used to generate simulations for how a meme might originate and travel through social media as it flows from user to user. In the process of developing a simulator using modern techniques for reducing complexity (such as object-oriented programming) and validating it by comparing its results with existing social media and big data, origins of memes can then be proposed and tested with the simulator to see if they are highly likely to be the origin.

In order to determine the utility of the SARF and M3D in guiding studies on hyperlocal human dynamics using social media and big data, more research is needed to further develop and test these frameworks. This position paper addresses potential opportunities and challenges for using frameworks like the SARF and M3D as analytical tools in this manner.

References

1. Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. *Wired Magazine 16*(7), 108-9.
2. Center for Human Dynamics in the Mobile Age. (2014, May 7). Retrieved from: http://humandynamics.sdsu.edu/
3. Faghmous, J. H., & Kumar, V. (2014). A big data guide to understanding climate change: The case for theory-guided data science. *Big Data*, *2*(3), 155-163.
4. Pidgeon, N., & Henwood, K. The social amplification of risk framework (SARF): Theory, critiques, and policy implications. In Bennett, P. Bennett (Ed.) *Risk communication and public health* (pp. 53-68). Oxford University Press.
5. Shah, D. V., Cappella, J. N., & Neuman, W. R. (2015). Big data, digital media, and computational social science possibilities and perils. *The ANNALS of the American Academy of Political and Social Science*, *659*(1), 6-13.
6. Spitzberg, B. H. (2014). Toward a model of meme diffusion (M3D). *Communication Theory*, *24*(3), 311-339.