

## Using Twitter as a Metric in Social Science and Public Health

By Papa Chakravarthy

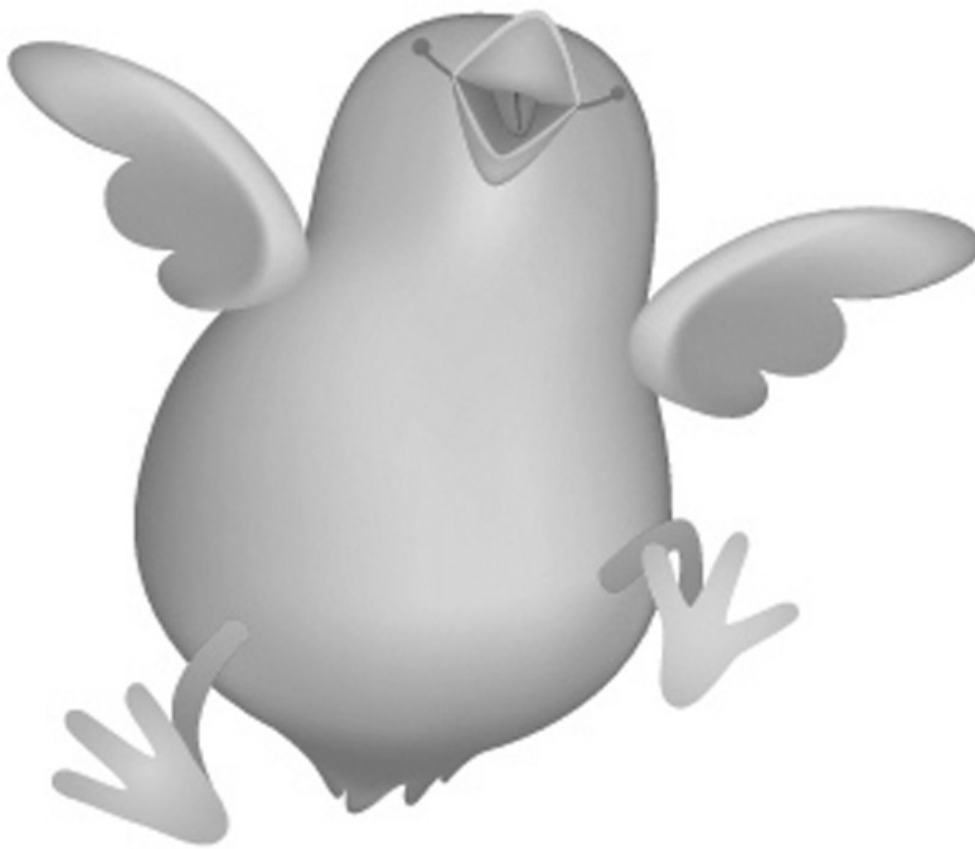
**Twitter is a** social network made up of individuals sharing their thoughts in short, 140 character phrases. Because Twitter is widespread, with over 200 million users posting their thoughts and comments in real time, it can be an extremely useful stream of information. The applications of Twitter are limited only by what people tweet, meaning that we can explore a variety of fields. However, before we can research populations through Twitter, we must understand the parameters of Twit-

ter's environment.

Even though there is a relatively short character limit on tweets, individually reading every tweet to decode them into topics would be impractical. Instead, researchers are relying heavily on models to make large numbers of tweets far easier to process. There are different ways to work with tweets but almost all of them rely on keyword analysis. Computer programs can analyze the words and phrases included in tweets and can give a category or mood

level of any given tweet. This strategy is limited by the ways people communicate, like the use of sarcasm. On the whole, however, when working with between ten and twenty billion tweets, these analyses are valuable.

Uses of Twitter in research are restricted by the same biases that accompany any research using social media. Although Twitter is a widely used application, it certainly does not provide complete coverage; for example, low income populations are only rarely represented. Twitter is also distinct from survey data because of the lack of an actual survey. People can tweet about whatever they like,



# twitter

meaning a tweet coded for “watching television” might actually be about how much a person hates watching television. Naturally, the computer programs coding tweets attempt to correct for problems like this, but they may not always be successful, muddling data and decreasing the significance of results. Moreover, the lack of a direct survey means Twitter research can pull information completely unrelated to the topic of interest. In order to find data in the numbers necessary to provide real clues about behavior, computers must process huge amounts of information. Despite these issues, Twitter research is a growing field and is currently making use of a few major tactics.

One possible application of this keyword analysis lies in global health. Keyword analysis can provide information about the subject of every single tweet. For example, a tweet like “Grey goose at da club!” would be categorized as being alcohol related. If many people tweet about alcohol in a given area, it would stand to reason that alcoholism rates are higher. Similar types of analyses can give information on obesity and health wellness in a variety of places. This type of subject analysis specified by Twitter’s geolocation information could be quite valuable.

There are other uses of Twitter, particularly when tracking individual users. By following the tweets of a certain person over time, we can find changes in behavior and link those changes to a cause. For example, if a keyword analysis shows a lot of alcohol related tweets abruptly stopping, perhaps

the individual stopped drinking. On the other hand, if a person exhibits that pattern and then alcohol tweets slowly reappear, perhaps alcoholism is a factor in behavior. This type of behavior analysis is slowly gaining steam in the world of Twitter research. The study of individuals can help people stay healthy and a similar type of study could work very well with regard to obesity.

Mood is another way of describing tweets. ANEW scores are a method of ranking mood that are popular in the psychiatric community and applying these scores to Twitter can have great potential. Studying the way people behave on social networks, like whether they only tweet happy thoughts, is decidedly interesting, but there are other uses. By tracking an individual user’s mood scores over time, we may be able to see distinct changes in mood. Studying these changes could lead to an understanding of how the human brain functions and might have relevance to the medical field. Current research is attempting to study depression through the lens of moods on Twitter. **H**

—*Papa Chakravarthy '12 is an Economics concentrator in Pforzheimer House.*

---

#### References:

1. <http://www.ccs.neu.edu/home/amislove/twitter-mood/>
2. <http://bits.blogs.nytimes.com/2010/07/22/tracking-u-s-mood-through-twitter/>
3. <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2816>