LANGER RESEARCH ASSOCIATES

SURVEY RESEARCH DESIGN • MANAGEMENT • ANALYSIS

Last updated: August 2013

Briefing paper: Social Media and Public Opinion

The use of social media to estimate public attitudes faces major challenges. From sampling to analysis, inherent difficulties in seeking valid and reliable estimates of public opinion from such data are apparent. And falsification, reportedly on a vast level, is a growing cause for concern.

Evaluating social media commentary may be potentially promising for trend-spotting, modeling or qualitative analysis, with development efforts under way in areas including economic research and health surveillance.[1] These, however, do not rise to the level of meaningful measurement of public attitudes or behavior, particularly within a calculable level of confidence. And even for impressionistic or qualitative analysis, evidence of extensive manipulation of postings via robotic computer programs, known as bots, is troublesome.

Beyond the reported presence of fabricated, orchestrated and/or automated postings, additional difficulties include the biasing nature of self-selection in participation in social media, including subject-specific use of the medium; challenges in deciphering often cryptic postings; sampling problems relating to keyword selection and variance in the volume of postings; and the lack of adequate or reliable demographic or attitudinal data with which to produce meaningful analysis.

Fundamentally, tweets, postings, clicks and "likes" do not reliably constitute individual-level commentary. One person can tweet, retweet or post a comment many times, and can use multiple identities in doing so. Such commentary, moreover, can be orchestrated in promotional or public relations campaigns by businesses, organizations, governments and computer hackers, via volunteers, paid agents or – increasingly – automated methods. And there are open questions about the effectiveness and possible impacts of efforts to filter out this noise.

Growing concerns have focused on the use of bots to produce ersatz social media content. A Forbes magazine article in August 2012, "Invasion of the Twitter Bots," cited a researcher's estimate that, in the 2012 presidential election, as many as 30 percent of Barack Obama's Twitter followers and 22 percent of Mitt Romney's were fabricated. Fast Company magazine reported that individuals can purchase 25,000 fabricated Twitter followers for $247, and that these "bots" will automatically tweet snippets of text.[2] The New York Times subsequently confirmed this account, and, in April 2013, reported that more than two dozen services sell fake Twitter

[1] Note, for example, Bloomberg News' coverage of efforts to use Google search records in models predicting retail sales and other consumer-driven trends, an approach said to be still in its infancy (Ito & Odenheimer, 2012).
[2] A year later, in August 2013, The New York Times reported the same: "Mercenary armies of bots can be bought on the Web for as little as $250."

followers, typically in batches of one thousand to one million accounts. It reported estimates that the most popular of these had produced as many as 20 million fake followers.

Subsequent news coverage has extended these findings. An analysis in The New York Times of Aug. 10, 2013, for example, included the following items:

- Researchers suggest that "only 35 percent of the average Twitter user's followers are real people" and that "within two years, about 10 percent of the activity occurring on social online networks will be masquerading bots."

- "Carina Santos," a much-followed Brazilian journalist on Twitter, ranked by Twitalyzer and Klout as having more online "influence" than Oprah Winfrey, was revealed in July 2013 not to be a real person, but rather a bot created by university researchers.

- Officials from Mexico's governing party were accused in 2012 of flooding Twitter with bogus posts designed to trip spam filters and shut down hashtags used by the party's critics. In Russia, thousands of Twitter bots posted hundreds of messages a day targeting anti-Kremlin activists in advance of 2011 parliamentary elections. The Syrian government has been accused of similar tactics.

- Before the 2010 midterm elections, per Indiana University researchers, two accounts sent out 20,000 similar tweets, most of them linking to or promoting the web site of then-House Minority Leader John Boehner. And in 2012, University of Illinois at Chicago researchers reported on the extensive use of bots to market e-cigarettes via social media.

Among previous reports, Forbes in 2012 cited complaints that large numbers of Facebook advertisement click-throughs may have been generated by automated bots. Other media cited a Facebook filing indicating that an estimated 8.7 percent of its user accounts were duplicates, misclassified or spam-related (Kelly, 2012). And The Economist, in an April 2013 report on microblogs and other social media in China, reported on "a flourishing black-market trade in fake 'followers' and 'retweets' to boost brands, celebrities and sometimes the microblog itself." It said: "Sina Weibo, for instance, has more than 500m registered accounts, but many of them are robots employed to generate artificial buzz. Sina itself says that the number of daily active users at the end of 2012 was only 46m."[3]

When actual individuals are in fact involved, their location – even their nationality – may not reliably be captured, leaving the universe ill-defined. And, even given real participants, the self-selected nature of who's tweeting, posting or clicking (and, as noted, how often) is inherently biasing. Even if it were possible to project to the population of all users of a particular social media site (as unlikely as that is), there is no theoretical basis on which to assume that the information that has been collected and evaluated is representative of the views of non-users.[4]

[3] In another example, the (London) Telegraph in April 2013 cited research suggesting that half of the pop star Justin Bieber's followers may be fake.
[4] A Pew Research Center survey in February 2012 found that 15 percent of online adults used Twitter, 8 percent of them daily, with usage rates ranging from 26 percent of online 18- to 29-year-olds to 14 percent of online 30- to 49-

Enormous variability in the number of daily posts or tweets on a single subject poses a further problem. There may be comparatively few for a substantial period of time, then a vast number in a very short period, often related to some event (e.g., news- or celebrity-based). Analysts are faced with deciding whether to allow a brief spike to dominate the data, or to smooth it; if so how, on what grounds and with what implications for time-trend analysis.

There's an additional, crucial difficulty in determining what posts or tweets mean. This requires content analysis, in which language is parsed, compared against a codebook and classified (e.g., as positive, negative or neutral). Content analysis traditionally is carried out by trained individuals, independent of the research hypothesis, working individually as coders; their work then is compared, with a very high level of intercoder reliability required – customarily at least 90 percent agreement.

Computerized sentiment analysis programs have been developed as a lower-cost and faster alternative to human coding of texts. Evaluations of commercially available products, however, indicate that they have significant problems, e.g., in parsing the slang, irony, sarcasm, synonyms, abbreviations, acronyms, emoticons, contextual meaning and hashtags that commonly are used online. Without a reliable assessment of the meaning of individual tweets or posts, there's no discernible way to construct a meaningful aggregate analysis of their content.

A related complication involves the use of word restrictions to select what text is included in a given analysis. An effort to measure posts or tweets on the president's popularity, for instance, might be limited for expediency to those including the word "Obama" or "president." Those choices, however, could create potentially biasing noncoverage by excluding other relevant words – "Barack," "the prez," "BO," "potus" and so forth. Misspellings and nonstandard constructions add further difficulties (e.g., "#barackobama").

Presenting at the annual conference of the American Association of Public Opinion Research in May 2012, Kim et al. compared validated human coding of tweets with three separate automated sentiment analysis systems, with highly inconsistent results. One automated system's classification of "positive" tweets on "health care reform" agreed with the human coding in a mere 3 percent of cases. Another system matched human coders in assessing "negative" tweets on another subject in 12 percent of cases.[5]

Many such mismatches were observed. Nor did the computerized systems agree in many cases with each other. As such, analysis of the meaning of these aggregated tweets would be based not on their actual content, but on the vagaries of the software used to categorize that content.

These challenges were echoed in a March 2013 article in AdAge.com, in which a Coca-Cola executive cited an internal 2010 study that found "widespread differences" in content analysis by human coders compared with a computerized analysis of more than 1,000 randomly selected

---

year-olds and 9 percent of online 50- to 64-year-olds. Another analysis suggests that less than 30 percent of Twitter's user base is American (Semiocast, 2012).

[5] Overall match rates ranged from 43 percent to 62 percent – roughly what chance would dictate – owing to a high match rate for "neutral" tweets, which are of less analytic interest.

social-media messages (especially in longer-format platforms such as Facebook and blogs, vs. shorter-format posts such as Twitter).[6] "When we say it's positive, the machine about 21 percent of the time says it's negative," Eric Schmidt, Coca-Cola's senior manager of marketing strategy and insights, was quoted as saying, "That can cause some problems in our understanding."

Schmidt also reportedly said the company found no measurable impact of online chatter on its sales. In what the article called "a stunning admission," he said that when "online buzz" was included in the company's standard analysis of the impact of digital media, "We didn't see any statistically significant relationship between our buzz and our short-term sales."[7]

Other papers have reported some success using experimental computer programs to discern the meaning of tweets (e.g., Agarwal et al., 2011; Pak & Paroubek, 2011). However these papers leave many questions open; we await documentation of a validated, commercially available product; and experimental efforts to use computer programs to derive meaning from tweets don't address the equally fundamental sampling issues cited above.

Additional challenges lie beyond difficulties in sampling and content analysis. One is the sparse nature of the material; tweets, for example, do not capture demographic and other attitudinal or behavioral information, which add invaluable context, depth and meaning to serious analysis of public opinion, quantitative and qualitative alike.[8] (While some social media websites do capture demographic data via user profiles, almost all demographic fields are optional, and access by researchers often is circumscribed. Researchers have reported experimental results in which demographic and attitudinal information can be imputed using publicly available information such as Facebook likes, Facebook or Twitter profiles and personal websites. Their accuracy aside, such efforts in the context of consumer marketing raise privacy concerns.[9])

Some researchers have sought to portray social media as representative of broader public opinion. In a 2010 conference paper, "From Tweets to Polls," O'Connor et al. correlated Twitter content analysis with results of random-sample public opinion surveys. The authors, however, concede "many ... issues with this text sample" and "clear opportunities for better preprocessing and stratified sampling," without identifying what those would be. They say that even "casual inspection" of their opinion estimate revealed "many examples of falsely detected sentiment." The correlations they report are highly variable based on smoothing periods, and disappear in certain time periods or with slight changes in keywords (e.g., "job" vs. "jobs"). And the paper offers no theoretical basis for why tweets might be generalizable to the full population.

Twitter itself created the Twitter Political Index, reputedly a representation of attitudes of its users on Obama and his Republican challenger, Mitt Romney, in 2012. A Twitter executive said

[6] See http://adage.com/article/cmo-strategy/coca-cola-sees-sales-impact-online-buzz-digital-display-effective-tv/240409/

[7] Schmidt said Coke was reviewing its definition of "buzz" and called the finding "by no means a generalized result that applies to all industries." Subsequently, a senior Coke marketing executive said social media played "a crucial role" in a broader marketing strategy, regardless of its direct impact on sales. See http://adage.com/article/cmo-strategy/social-media-matter-marketing-coca-cola/240444/

[8] See also a discussion of Google Surveys, below.

[9] See e.g. Kosinski et al. (2013).

the purpose was not to replace public opinion polling, but "to give journalists and researchers a more complete picture of the electorate" (Leydon, 2012). The scoring used in this index was opaque, but scores on Aug. 6, 2012, for instance, seemed to suggest that Obama was viewed three times as positively as Romney (at “61” vs. “24”), a result far out of line with contemporaneous probability-based surveys. Further, a day later, the numbers were “35” and “19,” indicating far more volatility than customarily is found in good-quality public opinion polls. Nor do publicly available explanations of the approach persuasively explain how it overcomes many of the challenges cited in this briefing paper (e.g., Little, 2012).

To the contrary, in a March 2013 report on a yearlong study, the Pew Research Center found that sentiment expressed on Twitter on politics and policy “often differs a great deal from public opinion as measured by surveys,” sometimes more liberal, sometimes more conservative and often more negative. “Twitter users are not representative of the general public,” Pew said, noting that they are, for example, younger and more likely to be Democrats than non-users. It suggested that this reflected the self-selected nature of participants, i.e., “both the narrow sliver of the public” using Twitter and the yet narrower slice and considerable variability in those who tweet on a particular topic.

Of the eight events studied, Twitter sentiment differed significantly from public sentiment in six cases. Some differences were dramatic: In one, Pew said that tweets responding to the first presidential debate in 2012 divided 59-40 percent in Obama’s favor, while polling results favored Romney’s performance by more than a 3-1 margin.

“Overall, the reaction to political events on Twitter reflects a combination of the unique profile of active Twitter users and the extent to which events engage different communities and draw the comments of active users,” the Pew report concluded. While this may be of interest, it added, “it does not reliably correlate with the overall reaction of adults nationwide.”

Some other elements raised in this paper were not addressed in the Pew study, including the possibility of orchestrated Twitter campaigns, fabricated accounts and the difficulty of selecting keywords to identify tweets for inclusion. (For example, its analysis of tweets relating to the 2013 State of the Union address was limited to those that included the words “state” and “union” or “Obama” or “SOTU,” with no assessment of how many relevant tweets this excluded.)

Contrary to Kim et al. (2012), Pew reported that the software it used to categorize Twitter content as positive or negative claims more than 90 percent reliability with human coders, and said its own check of the reliability “came up with similar results,” but no details of this analysis were provided.

Moving to Facebook, an article in Sociological Methods & Research (Bhutta, 2012) discusses an attempt to survey Catholics via “chain referral” (i.e. snowball) sampling on the social media site, said to be achievable “faster, cheaper, and with less assistance” than traditional methods. The author reports that participants “were disproportionately female, young, educated and religiously active,” but that their responses nonetheless “preserved key correlations” found in probability-sample data.

Some of the disproportions were impressive: Seventy percent of participants were female (vs. 53 percent of Catholic adults in probability-sample data from the General Social Survey), two-thirds held college degrees (vs. 24 percent in GSS data) and two-thirds reported attending Mass weekly (vs. 27 percent per GSS). In the Facebook data, more men than women reported attending Mass more than once a week, 28 vs. 22 percent, a result that "directly contradicts not only the GSS data but also a huge and varied body of research demonstrating greater religiosity among women than men for all aspects of religious behavior."

As such, and despite the correlations she goes on to examine, the author concludes, "the Facebook respondents cannot possibly serve as a representative sample of the general Catholic population. (Pollsters should view Facebook findings with extreme caution.)"

Other researchers have expressed similar concerns. In a paper titled "How Not to Predict Elections," Metaxas et al. (2011) report, "This research has revealed that data from social media did only slightly better than chance in predicting election results in the last U.S. congressional elections. We argue that this makes sense: So far, only a very rough estimation on the exact demographics of the people discussing elections in social media is known, while according to the state-of-the-art polling techniques, correct predictions requires [*sic*] the ability of sampling likely voters randomly and without bias." A path to approach better estimates, they suggest, "would be to establish a sampling method comparable to the ones used by professional pollsters, though there are many obstacles in doing so today."[10]

Google Surveys

While distinct from attempted analysis of social media postings, another effort to collect attitudinal data online has been developed by Google Consumer Surveys, which offers researchers the ability to present one or two survey questions to a subset of Google users trying to access "premium content" (such as news articles or videos). Results are supplemented with demographic data that have been imputed through analysis of the users' IP addresses and previous page views.

A self-published paper by Google introducing its approach apparently is used to support the marketing claim that Google surveys produce "results that are as accurate as probability based panels."[11] There is, however, no indication of why the population of selected Google premium content users would be representative of any broader population, nor any evidence to support such a claim. Questions on sampling and weighting include noncoverage of non-Google users (as well as non-premium content users), opt-out rates, adjustments for frequency of Google use and appropriate weighting parameters. Informal testing, further, indicates mismatches between

[10] Daniel Gayo-Ayello of the University of Oviedo, Spain, a co-author of the Metaxas paper, has published a summary of relevant papers from which he says "it can be concluded that the predictive power of Twitter regarding elections has been greatly exaggerated, and that hard research problems still lie ahead" (Gayo-Avello, 2012).
[11] See McDonald et al. (2012) and http://www.google.com/insights/consumersurveys/how.

Google's imputation and actual demographic data.[12] Beyond these issues, the limitation of asking only one or two survey questions would be inadequate for many research projects.[13]

A report by the Pew center comparing attitudes and demographic data between a dual frame telephone survey and Google Consumer Surveys had mixed results. Inferred demographic data were not available in 30 to 40 percent of cases, Pew reported, "either because their cookies are turned off but more often because the algorithm cannot determine a trend from the websites … that would suggest what gender or age they are." In cases in which inferences were made, moreover, 25 percent were misidentified as to their sex and 56 percent were misidentified as to their age category.

Attitudinal responses across 43 questions had median and mean sample-type differences of 3 and 6 percentage points, respectively; this, however, was "based on data that adjusted results for mode differences stemming from the presence or absence of explicit 'don't know' or 'no opinion' responses." Additionally, even with that adjustment, there were "several sizeable differences that ranged from 10 to 20 points." Support for same-sex marriage was 48 percent in the telephone survey but 59 percent in the Google survey; support for President Obama's immigration policy was 63 percent in the telephone survey vs. 52 percent in the Google survey; and belief that average temperatures have been rising globally was 67 percent vs. 57 percent. Unadjusted differences were not provided, nor were results of significance testing or a measure of the variability of mean differences needed to assess statistical significance of these differences overall. Also, the Google Surveys results were weighted to age, sex and state or region (as available) of all U.S. internet users, with no explanation of this choice (it not being a survey of that population).

Pew noted that limited multivariate data were available given Google's one- or two-question restriction, so similarities or differences between the Google and telephone survey data in the relationship among variables could not be assessed.

On coverage, Pew said, "It is unknown whether visitors to the network of publisher sites are fully representative of all internet users or what proportion of internet users are covered by the publisher network." Because probability of selection is unknown, "no meaningful margin of error can be calculated for projecting the results to the internet population," much less the general population. "In addition, the non-probability sampling may result in more variation from sample to sample."

Further research is welcome, and as noted, the use of social media for anecdotal purposes may show value. At this time, however, we see no persuasive evidence on which to accept compilations of social media postings, or other approaches described in this briefing paper, as valid and reliable measurements of public opinion.

---

[12] Using https://www.google.com/settings/ads/onweb/, one woman on our staff was misidentified as a 55-year-old male with an interest in beauty and fitness, another as a 65+ woman (double her age) with an interest in motorcycles and one man as a senior citizen in the Pacific Northwest with an interest in space technology.

[13] In addition to the one to two question limit, question stems are limited to 125 characters and response options to 44 characters. No more than five response categories are allowed.

See also our briefing paper on opt-in online panels:
http://www.langerresearch.com/uploads/Langer_Research_Briefing_Paper-Opt-in_Online_Panels.pdf

References

Akcora, C. G., Bayir, M. A., Demirbas, M., & Ferhatosmanoglu, H. (2010, July). *Identifying breakpoints in public opinion.* Paper presented at the 1st Workshop on Social Media Analytics, Washington, D.C.
http://upinion.cse.buffalo.edu/beta/SOMApaper.pdf

Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. (2011). Sentiment analysis of Twitter data. *Proceedings of the Workshop on Languages in Social Media* (pp. 30-38). Stroudsburg, PA: Association for Computational Linguistics.
http://www1.ccls.columbia.edu/~beck/pubs/lsm2011_full.pdf

Bialik, C. (2012, February 10). Mining tweets for public opinion. *The Wall Street Journal.* Retrieved from http://blogs.wsj.com/numbersguy/mining-tweets-for-public-opinion-1118/

Bhutta, C.B. (2012, March 21) Not by the book: Facebook as a sampling frame. *Sociological Methods & Research*, Sage Publications.
http://smr.sagepub.com/content/early/2012/03/19/0049124112440795.full.pdf+html

Bollen, J., Pepe, A., & Mao, H. (2011). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. *Proceedings of the 5th International Association for the Advancement of Artificial Intelligence Conference on Weblogs and Social Media* (pp. 450-453). Menlo Park, CA: The AAAI Press.
http://arxiv.org/pdf/0911.1583.pdf

Calzolari, M.C. (2012, August 8). Analysis of Twitter followers of US presidential election candidates: Barack Obama and Mitt Romney. Unpublished paper accessed at http://digitalevaluations.com/DigitalEvaluations-Obama_Romney.pdf

Cummings, D., Oh, H., & Wang, N. (2011). *Who needs polls? Gauging public opinion from Twitter data.* Unpublished manuscript.
http://nlp.stanford.edu/courses/cs224n/2011/reports/nwang6-davidjc-harukioh.pdf

Diakopoulos, N. A., & Shamma, D. A. (2010). Characterizing debate performance via aggregated Twitter sentiment. *Proceedings of the 28th International Conference on Human Factors in Computing Systems* (pp. 1195-1198). New York, NY: Association for Computing Machinery.
http://dmrussell.net/CHI2010/docs/p1195.pdf

Epstein, G. (2013, April 6) Microblogs: Small beginnings. *The Economist.* Retrieved from http://www.economist.com/news/special-report/21574632-microblogs-are-potentially-powerful-force-change-they-have-tread

Feifer, J. (2012, August 8). Who's that woman in the twitter bot profile? *Fast Company.* Retrieved from http://www.fastcompany.com/3000064/whos-woman-twitter-bot-profile

Finn, S., & Mustafaraj, E. (2012). Real-time filtering for pulsing public opinion in social media. *Proceedings of the 25th International Florida Artificial Intelligence Research Society Conference* (pp. 32-37). Menlo Park, CA: The AAAI Press. https://www.aaai.org/ocs/index.php/FLAIRS/FLAIRS12/paper/viewFile/4459/4762

Gayo-Avello, D., Metaxas, P. T., & Mustafaraj, E. (2011). Limits of electoral predictions using Twitter. *Proceedings of the 5th International Association for the Advancement of Artificial Intelligence Conference on Weblogs and Social Media* (pp. 490-493). Menlo Park, CA: The AAAI Press. http://cs.wellesley.edu/~pmetaxas/ICWSM11-limits_predict_elections.pdf

Gayo-Ayello, D. (2012). "I wanted to predict elections with Twitter and all I got was this lousy paper": A balanced survey on election prediction using Twitter data. Retrieved from http://arxiv.org/pdf/1204.6441v1.pdf

Golbeck, J. & Hansen, D. L. (2011). Computing political preference among Twitter followers. *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems* (pp. 1105-1108). New York, NY: Association for Computing Machinery. http://hcil2.cs.umd.edu/trs/2010-20/2010-20.pdf

Hill, K. (2012, August 9). Invasion of the Twitter bots. *Forbes.* Retrieved from http://www.forbes.com/sites/kashmirhill/2012/08/09/the-invasion-of-the-twitter-bots/

Ito, A., & Odenheimer, A. (2012, August 1). Your 119 billion Google searches now a Central Bank tool. *Bloomberg*. Retrieved from http://www.bloomberg.com/news/2012-08-02/your-119-billion-google-searches-now-a-central-bank-tool.html

Jackson, E. (2012, July 31). Why do some advertisers believe that 90% of Facebook ad clicks are from bots? *Forbes.* Retrieved from http://www.forbes.com/sites/ericjackson/2012/07/31/why-do-some-advertisers-believe-that-90-of-facebook-ad-clicks-are-from-bots/

Kelly, H. (2012, August 2). 83 million Facebook accounts are fakes and dupes. *CNN*. Retrieved from http://articles.cnn.com/2012-08-02/tech/tech_social-media_facebook-fake-accounts_1_facebook-accounts-facebook-profiles-facebook-estimates

Kim, A., Richards, A., Murphy J., Sage A., & Hansen H. (2012, May). *Can automated sentiment analysis of Twitter data replace human coding?* Paper presented at the 67th Annual conference of the American Association for Public Opinion Research, Phoenix, AZ.

Kouloumpis, E., Wilson, T., & Moore, J. (2011). Twitter sentiment analysis: The good, the bad and the OMG! *Proceedings of the 5th International Association for the Advancement of Artificial Intelligence Conference on Weblogs and Social Media* (pp. 538-541). Menlo Park, CA: The AAAI Press. http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2857/3251

Kosinski, M., Stillwell, D., and Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *PNAS Early Edition.* Retrieved from http://m.pnas.org/content/early/2013/03/06/1218772110.full.pdf

Leyden, C. (2012, August 1). Twitter debuts tweet-based supplement to polling. *CBS News.* Retrieved from http://www.cbsnews.com/8301-503544_162-57484746-503544/twitter-debuts-tweet-based-supplement-to-polling/

Lindsay, R. (2008, October 22). Predicting polls with Lexicon [Blog post]. Retrieved from http://languagewrong.tumblr.com/post/55722687/predicting-polls-with-lexicon

Linkins, J. (2012, January 5). Twitter: Should we get excited about its election predictions? (No.). *The Huffington Post.* Retrieved from http://www.huffingtonpost.com/2012/01/05/twitter-election-predictions_n_1187328.html

Little, M. (2012, August 1). 'Twindex' evaluates Obama's, Romney's Twitter performance. *Los Angeles Times.* Retrieved from http://articles.latimes.com/2012/aug/01/news/la-pn-twindex-evaluates-obama-romneys-twitter-performance-20120801

Lui, C., Metaxas, P. T., & Mustafaraj, E. (2011). On the predictability of the U.S. elections through search volume activity. *Proceedings of the International Association for Development of the Information Society International Conference on e-Society* (pp. 165-172). http://cs.wellesley.edu/~pmetaxas/e-Society-2011-GTrends-Predictions.pdf

McDonald, P., Mohebbi, M., & Slatkin, B. (2012). *Comparing Google Consumer Surveys to existing and non-probability based internet surveys* [White paper]. Retrieved from http://www.google.com/insights/consumersurveys/static/consumer_surveys_whitepaper.pdf

Mehta, R., Mehta, D., Chheda, D., Shah, C., & Chawan, P. M. (2012). Sentiment analysis and influence tracking using Twitter. *International Journal of Advanced Research in Computer Science and Electronics Engineering, 1*, 72-79.

Metaxas, P. T., Mustafaraj, E., & Gayo-Avello, D. (2011, October). *How (not) to predict elections.* Paper presented at the IEEE SocialCom Conference, Boston, MA. http://cs.wellesley.edu/~pmetaxas/How-Not-To-Predict-Elections.pdf

Neff, J. (2013). Buzzkill: Coca-Cola Finds No Sales Lift from Online Chatter, AdAge.com. http://adage.com/article/cmo-strategy/coca-cola-sees-sales-impact-online-buzz-digital-display-effective-tv/240409/

O'Connor, B. (2008, December 27). Facebook sentiment mining predicts presidential polls [Blog post]. Retrieved from http://brenocon.com/blog/2008/12/facebook-sentiment-mining-predicts-presidential-polls/

O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). From tweets to polls: Linking text sentiment to public opinion time series. *Proceedings of the 4th International Association for the Advancement of Artificial Intelligence Conference on Weblogs and Social Media* (pp. 122-129). Menlo Park, CA: The AAAI Press. http://www.cs.cmu.edu/~nasmith/papers/oconnor+balasubramanyan+routledge+smith.icwsm10.pdf

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. *Proceedings of the 7th International Conference on Language Resources and Evaluation* (pp. 1320-1326). Paris, France: European Language Resources Association. http://deepthoughtinc.com/wp-content/uploads/2011/01/Twitter-as-a-Corpus-for-Sentiment-Analysis-and-Opinion-Mining.pdf

Perloth, N. (2013, April 5) Fake Twitter followers become multimillion-dollar business. *The New York Times.* Retrieved from http://bits.blogs.nytimes.com/2013/04/05/fake-twitter-followers-becomes-multimillion-dollar-business/

Pew Research Center. (2013). Twitter reaction to events often at odds with overall public opinion. Retrieved from http://www.pewresearch.org/2013/03/04/twitter-reaction-to-events-often-at-odds-with-overall-public-opinion/

Pew Research Center for the People and the Press. (2012). A comparison of results from surveys by the Pew Research Center and Google Consumer Surveys. Retrieved from http://www.people-press.org/2012/11/07/a-comparison-of-results-from-surveys-by-the-pew-research-center-and-google-consumer-surveys/

Saez-Trumper, D., Meira, W., & Almeida, V. (2011). From total hits to unique visitors model for election's forecasting. *Proceedings of the Association for Computing Machinery Web Science 2011* (pp. 1-2). New York, NY: Association for Computing Machinery. http://www.websci11.org/fileadmin/websci/Posters/WebSci2011-Saez-Trumper.pdf

Semiocast. (2012, July 30). Twitter reaches half a billion accounts: More than 140 millions in the U.S. [Blog post]. Retrieved from http://semiocast.com/publications/2012_07_30_Twitter_reaches_half_a_billion_accounts_140m_in_the_US

Smith, A., & Brenner, J. (2012). Twitter use 2012. *Pew Internet & American Life Project.* Washington, DC: Pew Research Center. http://www.pewinternet.org/~/media//Files/Reports/2012/PIP_Twitter_Use_2012.pdf

Tumasjan, A., Sprenger, T. O, Sander, P. G., & Welpe, I. M. (2011). Election forecasts with Twitter: How 140 characters reflect the political landscape. *Social Science Computer Review, 29*, 402-418. http://ssc.sagepub.com/content/29/4/402.abstract

Urbina, I. (2013, Aug. 10) I Flirt and Tweet. Follow Me at #Socialbot. *The New York Times.* Retrieved from http://www.nytimes.com/2013/08/11/sunday-review/i-flirt-and-tweet-follow-me-at-socialbot.html

Wright, P. (2013, April 10) Half of Justin Bieber's Twitter followers 'fake.' *The Telegraph.* Retrieved from http://www.telegraph.co.uk/technology/twitter/9985936/Half-of-Justin-Biebers-Twitter-followers-fake.html