



Social media sway: Worries over political misinformation on Twitter attract scientists' attention

Author(s): Rachel Ehrenberg

Source: *Science News*, Vol. 182, No. 8 (OCTOBER 20, 2012), pp. 22-25

Published by: Society for Science & the Public

Stable URL: <http://www.jstor.org/stable/23351069>

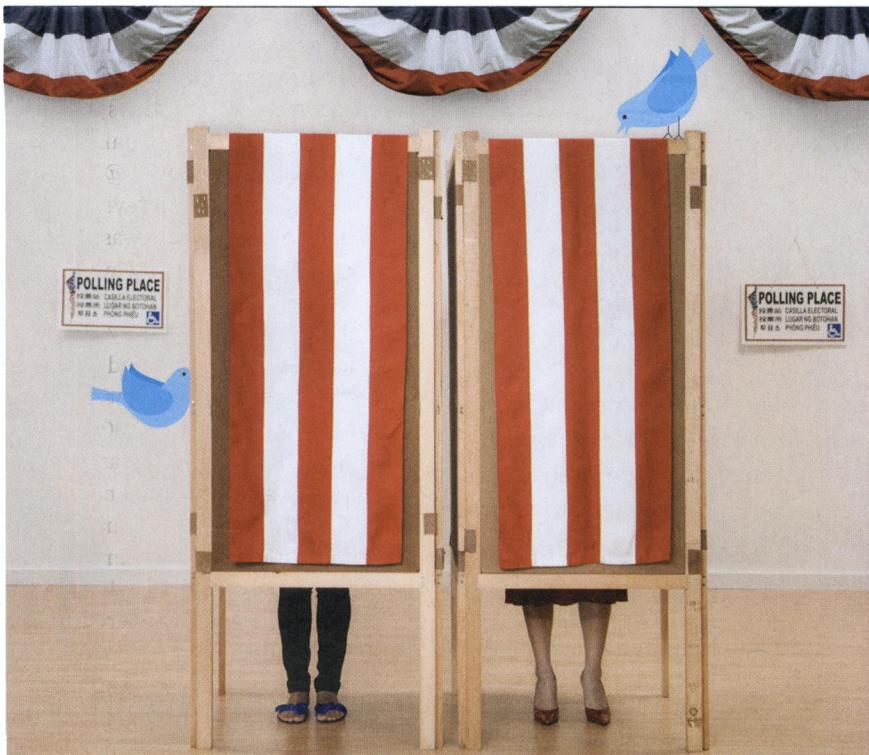
Accessed: 17-01-2018 20:40 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://about.jstor.org/terms>



Society for Science & the Public is collaborating with JSTOR to digitize, preserve and extend access to *Science News*



Social media sway

Worries over political misinformation on Twitter attract scientists' attention

By Rachel Ehrenberg

Four days before the 2010 special election in Massachusetts to fill the Senate seat formerly held by Ted Kennedy, an anonymous source delivered a blast of political spam. The smear campaign launched against Democratic candidate Martha Coakley quickly infiltrated the rest of the election-related chatter on the social networking service Twitter. Detonating over just 138 minutes, the “Twitter bomb” and the rancorous claims it brought with it eventually reached tens of thousands of people.

It's impossible to say whether the bomb left shrapnel that influenced the outcome of the heated race (Republican candidate Scott Brown overtook

Coakley in the campaign's final days). But the bomb did signal an end to the political left's dominance of social media. Twitter, which allows people to broadcast short online messages called “tweets,” has become a prominent player in the digital toolbox employed on both sides of the aisle. Campaigns and their supporters use the platform to spread messages, connect with like-minded people and garner votes. But along with shared news and engaging discussions come lies, propaganda and spin.

Though the strategic spread of misinformation is as old as elections themselves, the Internet Age has changed the game. Back before social media, the origins of political messages were less muddied. A man yelling on a soapbox looked like a man on a soapbox. Ads were ads. Most other material intended for wide consumption was vetted by journalists before it reached the masses. There were rumors and slander, of course, but those messages didn't get around so quickly.

Today venues such as Twitter offer a direct route for delivering a message to

a large target audience, often with little context for evaluating the message's veracity.

“Social media are a very effective and efficient way to spread false beliefs,” says political scientist Brendan Nyhan of Dartmouth College.

As pundits, journalists and citizens traverse the still-evolving social media landscape, scientists are doing the same. Using tools from linguistics, computer science and network science, these researchers are uncovering the digital calling cards of spin. Amid all the genuine discourse, teams are turning up speech dressed in truthful clothing squawked by impersonators, whether a single citizen with an agenda or a well-oiled political machine. While some may dismiss online misinformation as political graffiti that has no serious effect, others are concerned that it could change behavior at the voting booth.

Enough people are certainly paying attention. About 90 million Americans use Twitter in a typical month. Other social media sites are also well-populated: Estimates suggest that half of all Americans are on Facebook. And more than 5 million people in the United States spend part of their day on the blogging platform Tumblr.

While a hefty number of voters are on these platforms, no outlet is so mainstream that everybody uses it. The splintered nature of the social media landscape means misinformation often flies under the radar of the fact-checking apparatus employed by the traditional mass media, says Kathleen Hall Jamieson, director of the Annenberg Public Policy Center at the University of Pennsylvania.

“The danger is, correction channels aren't sitting inside that universe,” Jamieson says. “And people may vote for, or elect someone — or not vote — based on misinformation.”

Dirty digital tricks

The Coakley bomb, uncovered by computer scientists Eni Mustafaraj and Panagiotis Metaxas in the weeks after the 2010 special election, was cutting edge for its day.

PHOTO: HILL STREET STUDIOS; BIRDS: E. FELICIANO; BUNTING: DAY59/ISTOCKPHOTO

Back in 2006, the year Twitter was born, the going dirty digital political trick was to game Web search engines, says Mustafaraj, of Wellesley College in Massachusetts. During the 2006 midterm congressional elections, for example, the left-leaning group My Direct Democracy tried to manipulate search results to boost the prominence of negative stories about Republican incumbents. Through 2006, searching “miserable failure” on Google brought up President George W. Bush’s website.

By the 2008 congressional elections, such tactics became largely ineffective thanks to Google tweaking its search algorithms. An analysis by Mustafaraj and Metaxas found that during those elections, the top five Google results for queries about candidates consistently yielded the candidates’ official websites, their campaign websites and their Wikipedia entries. (Today, the first result of a Google search for “miserable failure” yields the Wikipedia entry defining “Google bomb.”)

There were a few exceptions. For the Republican 2008 senatorial candidate for Louisiana, some search results were still negative. A later analysis by Mustafaraj and Metaxas, presented at an MIT workshop last June, pinned the prominence of one of the pages on manipulation by liberal bloggers. But for the most part gaming search engines is a thing of the past.

Today, political noisemakers can direct an audience to websites through social media platforms. Such was the case with the Coakley Twitter bomb.

Using Twitter’s application programming interface, which allows researchers to collect and examine tweets, Mustafaraj

Following the candidates Social media has quickly become a staple in election campaigns. Below are some definitions and a comparison of how the 2012 presidential candidates stack up on Twitter.



SOURCE: TWITTER.COM, DATA TAKEN 10/2/12

Twitter accounts created in a 13-minute interval on January 15, four days before the election (account names included @CoakleySaidWhat, @CoakleySaidThat, @CoakleyAgainstU, @CoakleyAndU). Further research revealed that the coakleysaidit.com site was also created on January 15. On that day, over the course of 138 minutes, the nine Twitter accounts sent 929 tweets to 573 individuals. Those individuals passed along the messages (“retweeted”) to others, potentially reaching more than 60,000 people.

Twitter data doesn’t necessarily reveal a tweeter’s location or actual name, and the Wellesley scientists couldn’t pin the bomb on any single person. But well after the election, they discovered that the website was registered to the American Future Fund. This Iowa-based Republican-leaning group is known for its connection to the campaign against 2004 presidential nominee John Kerry that led to the term “swift boating,” now synonymous with smearing a politician with an untrue or unfair claim.

The Coakley Twitter bomb was an early case of what Filippo Menczer, a specialist in complex networks and Web data mining, calls “astroturfing.” To the untrained eye, a surge in vitriol against a candidate can appear to be a grassroots outcry, growing naturally from constituent concern or discontent. But in actuality, it’s machine-made artificial grass, or AstroTurf. Astroturfing campaigns (which are prohibited by Twitter policies) can give the impression that a discussion is truly representative of what a lot of people are thinking, Menczer says. This could prompt people to change their minds at the polls, he says, or to not vote.

and Metaxas hunted for news about the special election. In the days surrounding the election, they collected more than 234,697 tweets that contained the words “Coakley” or “Scott Brown.” Among tweets containing links to websites, a disproportionate number directed readers to “coakleysaidit.com.” This website, which appeared in 1,112 tweets, urged people to sign an online petition protesting Coakley’s “discrimination” against various groups.

Then the team discovered that the tweets containing the coakleysaidit.com link came from nine similarly named

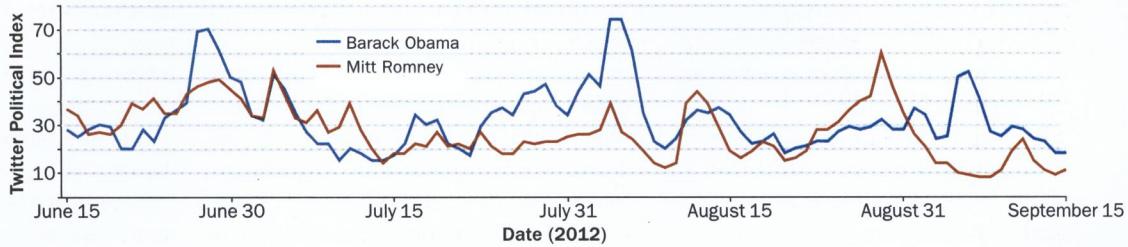
CLOCKWISE FROM TOP LEFT: © 2008 PETE SOUZA; GAGE SKIDMORE/WIKIMEDIA COMMONS; E. FELICIANO

Next-gen polling

A website called the Twitter Political Index, or Twindex, offers a daily appraisal of the sentiment surrounding each candidate on Twitter. The site’s algorithm analyzes the content of three days’ worth of tweets and then compares results with that day’s tweets about the candidates.

SOURCE: ELECTION.TWITTER.COM

	June 28	August 10	August 27–30	September 4–6	September 12
	Supreme Court upholds President Obama’s health care law	Romney names Paul Ryan as his running mate	Republican National Convention meets and nominates Romney	Democratic National Convention convenes	White House confirms death of Libyan ambassador



Truthy tracks deceit

Menczer, of Indiana University Bloomington, heard about the Twitter bomb research at the 2010 Web Science Conference in Raleigh, N.C. He figured there were probably many more instances of such social media skullduggery, so he began a project to track the spread of ideas and phrases on Twitter.

Called “Truthy,” the project captures thousands of tweets per hour and searches for tweets associated with particular topics by looking for hashtags (the # symbol) embedded in the 140-character-or-less messages. When placed in front of a word or phrase on Twitter, the # sign tags the tweet as having a particular

Political bomb After the 2010 special election in Massachusetts, scientists discovered nine similarly named Twitter accounts created within minutes of each other on January 15 (below), four days before the election. In a little over two hours, the suspicious accounts sent out 929 tweets.

Accounts created for spam attack

Account Name	Time created (1/15/10)	No. of tweets
@CoakleySaidWhat	18:43:46	28
@CoakleyWhat	18:44:55	127
@CoakleySaidThat	18:46:12	125
@CoakleyAgainstU	18:48:21	127
@CoakleyCatholic	18:50:22	127
@CoakleyER	18:52:05	127
@CoakleyAG	18:53:17	32
@CoakleyMass	18:54:31	109
@CoakleyAndU	18:56:02	127

Tweets from spamming accounts

Links in the suspicious tweets (examples below) sent readers to an online petition against Democratic candidate Martha Coakley, created the same day as the suspicious Twitter accounts.

AG Coakley thinks Catholics shouldn't be in the ER, take action now! bit.ly/8gDSp5

Tell AG Coakley not to discriminate against Catholics in medicine! bit.ly/8gDSp5

Catholics can practice medicine too! Tell AG Coakley today. bit.ly/7yXbTd

Sign the petition to AG Coakley today. We won't tolerate discrimination of any kind! bit.ly/8gDSp5

relevance. For example, #tcot identifies tweets referencing “top conservatives on Twitter.” The researchers (and anyone who goes to the Truthy site) can also search for a particular phrase or tweets sent by, sent to or sent about a particular tweeter, as indicated by the @ that appears as part of an account’s username, as in @BarackObama.

With network analysis techniques that diagram relationships among tweeters, those who follow them (that is, receive their tweets) and those who choose to retweet messages, Menczer and his colleagues can trace where messages originate and how they travel. Tools from linguistics that evaluate words and phrases also allow a rough gauge of the political partisanship and the positive or negative sentiment surrounding a topic or tweeter.

Via the Truthy system, the researchers have uncovered several examples of astroturfing. Leading up to the 2010 midterm elections, for example, two suspicious accounts were created within 10 minutes of each other. @PeaceKaren_25 generated more than 10,000 tweets in a few months, almost all of which included positive messages about Republican candidates. The other, @HopeMarie_25, retweeted all the tweets generated by @PeaceKaren_25 but never created any original tweets. Neither account revealed its creator’s identity, and both have since been shut down, says Menczer, who presented an update on the Truthy system in Vancouver at a February meeting of the American Association for the Advancement of Science.

Menczer argues that generating Web traffic about a candidate or political platform lends credibility to the message, whether it is true or not. Tweets by prolific and influential tweeters may come up in search engine results, further suggesting that the messages are part of a widespread discussion.

And efforts to correct misinformation frequently fail. Evidence suggests that corrections on Twitter don’t always have the same wings as an original false claim.

“Misinformation can win out, even

with a correction,” says Jamieson of the Annenberg Public Policy Center, which runs a nonpartisan website called FactCheck.org that aims to call out inaccurate, misleading or false claims by politicians. “It depends on the frequency of the correction compared to the frequency of the misinformation.”

Menczer suspects that the accounts @PeaceKaren_25 and @HopeMarie_25 were “bots” designed to send out spam. These spam spreaders may be automated accounts, or duplicate accounts run by the same person or, by some definitions, any account that is intended to mislead. Spam bots, political or otherwise, are prohibited by Twitter, which has its own algorithms that monitor for such deceit. This type of monitoring includes looking for accounts that tweet misleading links. Other efforts focus on how people react to the account — how many people have blocked it, for example.

After Republican presidential candidate Mitt Romney’s Twitter account acquired 141,000 followers in two days in July, many suspected the new followers were fake. Devin Gaffney and Alexander Furnas, then at the University of Oxford’s Internet Institute in England, decided to investigate the followers by looking at the followers’ followers.

“A lot of the existing ways to detect cheaters focus on the actors themselves,” Gaffney says. “We’re asking, how do people react to this account?”

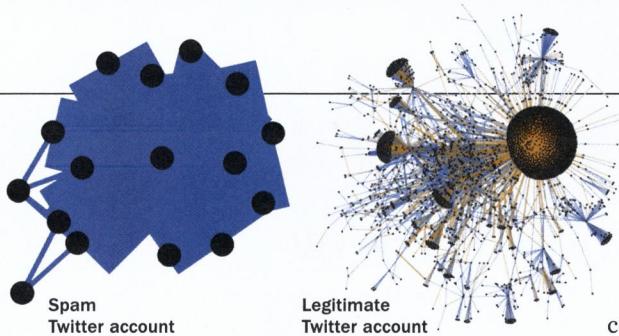
The approach leveraged a version of Alan Turing’s famous test for artificial intelligence: Do other people find Romney’s followers believable as human beings? After comparing a random sample of accounts that follow Romney’s followers with those of similarly sized Twitter accounts, Romney’s followers looked especially fishy. Of the followers of the randomly sampled accounts, about 10 percent had fewer than two followers. Yet nearly 27 percent of Romney’s newly acquired followers had just one or zero followers.

“No one follows them,” says Gaffney, who with Furnas describes the analysis in a piece in *The Atlantic*. “Clearly, they don’t pass the Turing test.”

SOURCE: P.T. METAXAS AND E. MUSTAFARAJ/2010

Moving messages

A diagram reveals a classic spam network (left), in which one account (center dot) promotes a website and other spam accounts collude in its spread. The pattern differs for mentions and tweets by a legitimate account, @sarhaplinusa (right).



While some speculate the Romney followers were created to make him appear more popular, others suggest that they were designed to make him look like a buffoon who has to buy Twitter followers – which is also against Twitter's rules.

Bots in charge

Going forward, fake accounts may become harder to detect. There are new levels of sophistication, says Tim Hwang of the San Francisco-based Pacific Social Architecting Corp.

Hwang should know; he's been developing virtual robots that are so real that they can change the shape of a social network, creating ties between people who weren't previously connected.

Last year, Hwang and his colleagues at the Web Ecology Project organized a competition to see who could build the best network-influencing social robots. Three teams had two weeks to write the code – the “brains” behind the bots – and then the bots had to infiltrate a network of 500 Twitter users. Each bot received one point for each target user that ended up following it and three points for each tweet sent by a target user that mentioned the bot. A team lost 15 points if its bot was recognized and shut down. By the end of the two-week competition, the creators of “James M. Titus” had won, with 107 followers and 198 mentions, for a total of 701 points.

Research suggests that by 2015, about 10 percent of a given person's social online network will be masquerading bots, presumably not obvious spam bots, but the social kind. Hwang and colleagues Ian Pearce and Max Nanis now create bots for clients they will not name. Instead of just having fake accounts pepper people with tweets, the

PacSocial bots have sleep-wake cycles so they appear more real. Some bots have built-in databases of current events, so they can piece together phrases that seem relevant.

Marketing products with these bots is an obvious task: Nanis describes a social bot that might post a Twitter picture of the African savanna with the tweet “this Nikon camera really was the best choice! Loving my trip.”

Bots that create chatter to try to sway elections are another obvious choice. Whether such bots will be detectable – by Twitter or by researchers monitoring the service – remains to be seen.

Because of the difficulty in correcting misinformation once it is out there, researchers are now trying to track rumors as they develop, rather than after they've fully infiltrated the Twitter-sphere. With the help of a \$2 million grant from DARPA, Menczer and others are building the machinery and algorithms to do real-time detection on large-scale datasets.

Dartmouth's Nyhan is also working

with colleagues including information scientist Paul Resnick of the University of Michigan in Ann Arbor on a project called “Fact Spreaders” that aims to recruit people in social networking communities to help spread accurate information to counter false claims.

It's hard to estimate the effect that ongoing digital discussions – true or false – could have on elections. In a recent study, Facebook users were more likely to go to the polls when they saw a message including pictures and names of friends who had voted (see Page 12).

Figuring out what type of information from whom yields the most power and when may help scientists beat misinformation to the punch. Swaying power could be most effective in influencing smaller elections, for which people don't necessarily have a lot of exposure to candidates' platforms.

“I don't expect a given rumor about a presidential candidate to dramatically influence the outcome of an election,” Nyhan says. “But false claims crowd out the more informed and fact-based debate we could be having,” he says. “It pollutes the political debate of our democracy.”

Explore more

- Truthy research at Indiana University: Truthy.indiana.edu

False news travels fast

Misinformation can spread quickly through social media, but corrections may not travel as far or fast. Last November, two days after Occupy Wall Street protesters were evicted from Zuccotti Park in New York City, the local NBC station sent out a tweet that the New York Police Department was closing air space above the protests, a message quickly retweeted by others. A few minutes later, NYPD tweeted that the information was incorrect. While NBC immediately sent out corrections, an analysis by Gilad Lotan of the social analytics company SocialFlow reveals that the correction (blue) wasn't tweeted as much as the incorrect claim (green).

Spread of @NBCNews/@NBCNewYork tweets, November 17, 2011

