Chapter 3

Can Technology Save Us?

Technology of Fake News

Fake news sites target the filter bubbles of groups most aligned with that news. They use the power of social media to do so. Initially fake news of the social media era was relatively easy to spot. The claims of early social media fake news purveyors were often meant as entertainment. Language, fonts, and links were often indicators that could be used to determine veracity. It took only a short time for fake news to become more insidious, more plentiful, more subtle, and subverted for manipulation of information and public opinion. Fake news has many new social media outlets where it can appear and can spread quickly via both human and nonhuman actors. During the 2016 presidential election cycle for example, fake news appeared often.[1] Determining what news was to be believed and what news was to be ignored became more a case of party affiliation than good sense.

Fake news sites and stories are shared for many different reasons. Some readers find the stories amusing. Some find them alarming. Others find them affirming of their beliefs. Many people share fake news without ever having read the content of the article.[2] Sharing of fake news, whether because it is amusing or because people think it is real, only exaggerates the problem. Did Pope Francis endorse candidate Donald Trump? No, but that didn't stop the story from appearing on social media and spreading widely.[3] Did Hillary Clinton run a child sex ring out of a Washington, DC, pizza shop? No, but that didn't stop a man with a gun from going there to exact vengeance.[4]

In the early days of the internet, fake news was not a big problem. There were some websites that sought to spoof, mislead, or hoax, but mostly it was all in good fun. While some websites sought to spread misinformation, their numbers were limited. It seemed as if the authority to shut down malicious websites was invoked more often. Creating a website on the early internet took time, effort, and computer programming skills that limited the number of people who could create fake news sites.

During the last decade, as an offshoot of the stream of information provided by the internet, social media platforms, such as Facebook and MySpace, were invented so that individuals could connect with others on the internet to point them to websites, share comments, describe events, and so on.

Following that came the invention of another type of social media—Twitter—which allows people to send very brief messages, usually about current events, to others who choose to receive those messages. One could choose to "follow" former President Barak Obama's Twitter postings—to know where he is going, what is on his agenda, or what is happening at an event. This kind of information can be very useful for getting on-site information as it happens. It has proved useful in emergency situations as well. For example, during the Arab Spring uprisings, Twitter communications provided information in real time as events unfolded.[5] During Hurricane Sandy, people were able to get localized and specific information about the storm as it happened.[6] Twitter is also a convenient means of socializing, for getting directions, and for keeping up-to-date on the activities of friends and family.

The power of the various tools that use the power of the internet and the information supplied there is epic. The spread of the technology required to make use of these tools has been rapid and global. As with most tools, the power of the internet can be used for both good and evil. In the last decade, the use of the

internet to manipulate, manage, and mislead has had a massive upswing.

Big Data

The collection of massive amounts of data using bots has generated a new field of study known as "big data."[7] Some big data research applies to the activities of people who use the internet and social media. By gathering and analyzing large amounts of data about how people use the internet, how they use social media, what items they like and share, and how many people overall click on a link, advertisers, web developers, and schemers can identify what appear to be big trends. Researchers are concerned that big data can hide biases that are not necessarily evident in the data collected, and the trends identified may or may not be accurate.[8] The use of big data about social media and internet use can result in faulty assumptions and create false impressions about what groups or people do or do not like. Manipulators of big data can "nudge" people to influence their actions based on the big data they have collected.[9] They can use the data collected to create bots designed to influence populations.[10]

Bots

Information-collecting capabilities made possible by harnessing computer power to collect and analyze massive amounts of data are used by institutions, advertisers, pollsters, and politicians. Bots that collect the information are essentially pieces of computer code that can be used to automatically respond when given the right stimulus. For example, a bot can be programmed to search the internet to find particular words or groups of words. When the bot finds the word or words it is looking for, its programming makes note of the location of those words and does something with them. Using bots speeds up the process of finding and collecting sites that have the required information. The use of bots to collect data and to send data to specific places allows research to progress in many fields. They automate tedious and time-consuming processes, freeing researchers to work on other tasks.

Automated programming does good things for technology. There are four main jobs that bots do: "Good" bots crawl the web and find website content to send to mobile and web applications and display to users. They search for information that allows ranking decisions to be made by search engines. Where use of data has been authorized, the data is collected by bot "crawlers" to supply information to marketers. Monitoring bots can follow website availability and monitor the proper functioning of online features.

This kind of data collection is useful to those who want to know how many people have looked at the information they have provided. "In 1994, a former direct mail marketer called Ken McCarthy came up with the clickthrough as the measure of ad performance on the web. The click's natural dominance built huge companies like Google and promised a whole new world for advertising where ads could be directly tied to consumer action."[11] Counting clicks is a relatively easy way to assess how many people have visited a website. However, counting clicks has become one of the features of social media that determines how popular or important a topic is. Featuring and repeating those topics based solely on click counts is one reason that bots are able to manipulate what is perceived as popular or important. Bots can disseminate information to large numbers of people. Human interaction with any piece of information is usually very brief before a person passes that information along to others. The number of shares results in large numbers of clicks, which pushes the bot-supplied information into the "trending" category even if the information is untrue or inaccurate. Information that is trending is considered important.

Good bots coexist in the technical world with "bad" bots. Bad bots are not used for benign purposes, but rather to spam, to mine users' data, or to manipulate public opinion. This process makes it possible for bots to harm, misinform, and extort. The *Imperva Incapsula* "2016 Bot Traffic Report" states that approximately 30 percent of traffic on the internet is from bad bots. Further, out of the 100,000 domains that were studied for the report, 94.2 percent experienced at least one bot attack over the ninety-day period of the study.[12] Why are bad bots designed, programmed, and set in motion? "There exist entities with both strong motivation and technical means to abuse online social networks—from individuals aiming to artificially boost their popularity, to organizations with an agenda to influence public opinion. It is not difficult to automatically target particular user groups and promote specific content or views. Reliance on social media may therefore make us vulnerable to manipulation."[13]

In social media, bots are used to collect information that might be of interest to a user. The bot crawls the internet for information that is similar to what an individual has seen before. That information can then be disseminated to the user who might be interested. By using keywords and hashtags, a website can attract bots searching for specific information. Unfortunately, the bot is not interested in the truth or falsehood of the information itself.

Some social bots are computer algorithms that "automatically produce content and interact with humans on social media, trying to emulate and possibly alter their behavior. Social bots can use spam malware, misinformation slander or even just noise"

to influence and annoy.[14] Political bots are social bots with political motivations. They have been used to artificially inflate support for a candidate by sending out information that promotes a particular candidate or disparages the candidate of the opposite party. They have been used to spread conspiracy theories, propaganda, and false information. Astroturfing is a practice where bots create the impression of a grassroots movement supporting or opposing something where none exists. Smoke screening is created when a bot or botnet sends irrelevant links to a specific hashtag so that followers are inundated with irrelevant information.

> When disguised as people, bots propagate negative messages that may seem to come from friends, family or people in your crypto-clan. Bots distort issues or push negative images of political candidates in order to influence public opinion. They go beyond the ethical boundaries of political polling by bombarding voters with distorted or even false statements in an effort to manufacture negative attitudes. By definition, political actors do advocacy and canvassing of some kind or other. But this should not be misrepresented to the public as engagement and conversation. Bots are this century's version of push polling, and may be even worse for society.[15]

Social bots have become increasingly sophisticated, such that it is difficult to distinguish a bot from a human. In 2014, Twitter revealed in a SEC filing that approximately 8.5 percent of all its users were bots, and that number may have increased to as much as 15 percent in 2017.[16] Humans who don't know that the entity sending them information is a bot may easily be supplied with false information.

Experiments in Fake News Detection

Researchers have studied how well humans can detect lies. Bond and DePaulo analyzed the results of more than 200 lie detection experiments and found that humans can detect lies in text only slightly better than by random chance.[17] This means that if a bot supplies a social media user with false information, that person has just a little better than a 50 percent chance of identifying the information as false. In addition, because some bots have presented themselves and been accepted by humans as "friends," they become trusted sources, making the detection of a lie even more difficult.

To improve the odds of identifying false information, computer experts have been working on multiple approaches to the computerized automatic recognition of true and false information.[18]

Written Text

Written text presents a unique set of problems for the detection of lies. While structured text like insurance claim forms use limited and mostly known language, unstructured text like that found on the web has an almost unlimited language domain that can be used in a wide variety of contexts. This presents a challenge when looking for ways to automate lie detection. Two approaches have been used recently to identify fake news in unstructured text. Linguistic approaches look at the word patterns and word choices, and network approaches look at network information, such as the location from which the message was sent, speed of response, and so on.[19]

Linguistic Approaches to the Identification of Fake News

The following four linguistic approaches are being tested by researchers:

In the Bag of Words approach, each word in a sentence or paragraph or article is considered as a separate unit with equal importance when compared to every other word. Frequencies of individual words and identified multiword phrases are counted and analyzed. Part of speech, location-based words, and counts of the use of pronouns, conjunctions, and negative emotion words are all considered. The analysis can reveal patterns of word use. Certain patterns can reliably indicate that information is untrue. For example, deceptive writers tend to use verbs and personal pronouns more often, and truthful writers tend to use more nouns, adjectives, and prepositions.[20]

In the Deep Syntax approach, language structure is analyzed by using a set of rules to rewrite sentences to describe syntax structures. For example, noun and verb phrases are identified in the rewritten sentences. The number of identified syntactic structures of each kind compared to known syntax patterns for lies can lead to a probability rating for veracity.[21]

In the Semantic Analysis approach, actual experience of something is compared with something written about the same topic. Comparing written text from a number of authors about an event or experience and creating a compatibility score from the comparison can show anomalies that indicate falsehood. If one writer says the room was painted blue while three others say it was painted green, there is a chance that the first writer is providing false information.[22]

In Rhetorical Structure (RST), the analytic framework identifies relationships between linguistic elements of text. Those comparisons can be plotted on a graph, Vector Space Modeling (VSM) showing how close to the truth they fall.[23]

Networks

In approaches that use network information, human classifiers identify instances of words or phrases that are indicators of deception. Known instances of words used to deceive are compiled to create a database. Databases of known facts are also created from various trusted sources.[24] Examples from a constructed database of deceptive words or verified facts can be compared to new writing. Emotion-laden content can also be measured, helping to separate feeling from facts. By linking these databases, existing knowledge networks can be compared to information offered in new text. Disagreements between established knowledge and new writing can point to deception.[25]

Social Network Behavior using multiple reference points can help social media platform owners to identify fake news.[26] Author authentication can be verified from internet metadata.[27] Location coordination for messages can be used to indicate personal knowledge of an event. Inclusion or exclusion of hyperlinks is also demonstrative of trustworthy or untrustworthy sources. (For example, TweetCred, available as a browser plugin, is software that assigns a score for credibility to tweets in real time, based on characteristics of a tweet such as content, characteristics of the author, and external URLs.[28]) The presence or absence of images, the total number of images by multiple sources, and their relationships and relevance to the text of a message can also be compared with known norms and are an indicator of the truth of the message. Ironically, all of this information can be collected by bots.

Experiments in Bot and Botnet Detection

A variety of experiments have been conducted using multiple processes to create a score for information credibility.[29] Research groups are prepared to supply researchers with data harvested from social media sites. Indiana University has launched a project called Truthy.[30] As part of that project, researchers have developed an "Observatory of Social Media." They have captured data about millions of Twitter messages and make that information available along with their analytical tools for those who wish to do research. Their system compares Twitter accounts with dozens of known characteristics of bots collected in the Truthy database to help identify bots.

Truthy
http://truthy.indiana.edu/about/

DARPA, Defense Advanced Research Projects Agency, is a part of the US Department of Defense. It is responsible for the development of emerging technologies that can be used by the US military. In early 2015, DARPA sponsored a competition whose goal was to identify bots known as influence bots. These bots are "realistic, automated identities that illicitly shape discussions on social media sites like Twitter and Facebook, posing a risk to freedom of expression."[31] If a means of identifying these bots could be discovered, it would be possible to disable them. The outcome of the challenge was that a semi-automated process that combines inconsistency detection and behavioral modeling, text analysis, network analysis, and machine learning would be the most effective means of identifying influence bots. Human judgment added to the computer processes provided the best results.

Many other experiments in the identification of bots have been reported in the computer science literature.[32] Bots and botnets often have a specific task to complete. Once that task is completed, their accounts are eliminated. Detecting bots and botnets before they can do harm is critical to shutting them down. Unfortunately, the means for detecting and shutting down bots are in their infancy. There are too many bot-driven accounts and too few means for eliminating them.

What happens to the information that bots collect is one part of the story of fake news. During the 2016 US presidential campaign, the internet was used to advertise for political candidates. Official campaign information was created by members of each politician's election team. News media reported about candidates' appearances, rallies, and debates, creating more information. Individuals who attended events used social media to share information with their friends and followers. Some reports were factual and without bias. However, because political campaigns involve many people who prefer one candidate over another, some information presented a bias in favor of one candidate or not favoring another candidate.

Because it is possible for anyone to launch a website and publish a story, some information about the political candidates was not created by any official of the campaign. In fact, many stories appeared about candidates that were biased, taken out of context, or outright false. Some stories were meant as spoof or satire; others were meant to mislead and misinform. One story reported that the pope had endorsed presidential candidate Donald Trump. In any other context, the reader would likely have no trouble realizing that this story was not true.

Enter the bots. There have been some alarming changes in how, where, and for what bots are used in the past ten years. Bots are being programmed to collect information from social media accounts and push information to those accounts that meet certain criteria.

> Social networks allow "atoms" of propaganda to be directly targeted at users who are more likely to accept and share a particular message. Once they inadvertently share a misleading or fabricated article, image video or meme, the next person who sees it in their social feed probably trusts the original poster, and goes on to share it themselves. These "atoms" then rocket through the information ecosystem at high speed powered by trusted peer-to-peer networks.[33]

Political bots have been central to the spread of political disinformation. According to Woolley and Guilbeault, the political bots used in the 2016 US elections were primarily used to create manufactured consensus:

> Social media bots manufacture consensus by artificially amplifying traffic around a political candidate or issue. Armies of bots built to follow, retweet, or like a candidate's content make that candidate seem more legitimate, more widely supported, than they actually are. Since bots are indistinguishable from real people to the average Twitter or Facebook user, any number of bots can be counted as supporters of candidates or ideas. This theoretically has the effect of galvanizing political support where this might not previously have happened. To put it simply: the illusion of online support for a candidate can spur actual support through a bandwagon effect.[34]

The Computational Propaganda Research project has studied the use of political bots in nine countries around the world. In Woolley and Guilbeault's report on the United States, the authors state, "Bots infiltrated the core of the political discussion over Twitter, where they were capable of disseminating propaganda at mass-scale. Bots also reached positions of high betweenness centrality, where they played a powerful role in determining the flow of information among users.[35]

Social bots can affect the social identity people create for themselves online. Bots can persuade and influence to mold human identity.[36] Guilbeault argues that online platforms are the best place to make changes that can help users form and maintain their online identity without input from nonhuman actors. To do that, researchers must identify and modify features that weaken user security. He identifies four areas where bots infiltrate social media:

1. Users create profiles to identify themselves on a social media platform. It is easy for bots to be programmed to provide false information to create a profile. In addition, the accessibility of the information in the profiles of other social media users is relatively easy to use to target specific populations.
2. In person, humans rely of a wide range of signals to help determine whether or not they want to trust someone. Online users have more limited options, making it much easier for bots to pretend to be real people. For platforms like Twitter, it is significantly easier to imitate a human because the text length is short and misspellings, bad grammar, and poor syntax are not unusual. Guilbeault indicates that popularity scores are problematic. He suggests, for example, "making popularity scores optional, private, or even nonexistent may significantly strengthen user resistance to bot attacks."[37]
3. People pay attention to their popularity in social media. A large number of friends or followers is often considered to be a mark of popularity. That can lead to indiscriminate acceptance of friend requests from unknown individuals, providing a place for social bots to gain a foothold. Bots send out friend requests to large numbers of people, collect a large following, and, as a result, become influential and credible in their friend group.
4. The use of tools such as emoticons and like buttons help to boost the influence of any posting. Bots can use the collection of likes and emoticons to spread to other groups of users. This process can eventually influence topics that are trending on Twitter, creating a false impression of what topics people are most interested at a given time. This can, of course, deflect interest in other topics.[38]

While Guilbeault has identified practices on social media platforms where improvements or changes could be made to better protect users, those changes have yet to be made. A groundswell of opinion is needed to get the attention of social media platform makers. The will to remove or change a popular feature such as popularity rating doesn't seem likely in the near future. In fact, while research is being done in earnest to combat the automated spread of fake or malicious news, it is mostly experimental in nature.[39] Possible solutions are being tested, but most automatic fake news identification software is in its infancy. The results are promising in some cases, but wide application over social media platforms is nowhere in sight. The research that exists is mostly based on identifying and eliminating accounts that can be shown to be bots. However, by the time that has been accomplished, whatever the bot has been programmed to do has already been done. There are very few means to automatically identify bots and botnets and disable them before they complete a malicious task.

Google and Facebook Anti–Fake News Efforts

The social media platforms and search engines themselves have made some efforts to help detect and flag fake news. Facebook created an "immune system" to

help protect itself from infection by bots.[40] Google announced that it will increase its regulation of advertising and linked-to websites.[41] Facebook has turned over the verification of information to five leading fact-checking organizations.[42] Facebook has also initiated a feature in parts of Europe called Related Articles, which provides readers with access to the results of fact-checking of original stories.[43] Google Digital News Initiative is creating programs to help users verify information themselves with Factmata. Overall, these attempts are reactive at best. The sheer volume of potential misinformation and the difficulty in identifying and shutting down bot accounts make these attempts seem feeble.

Factmata
http://factmata.com/

It seems that the battle of the computer programmers will continue indefinitely. When one side develops a new means of manipulating information to mislead, misinform, or unduly influence people, the other side finds a way to counter or at least slow the ability to make use of the new idea. This cycle continues in a seemingly endless loop. Using technology to identify and stop fake news is a defensive game. There does not appear to be a proactive means of eliminating fake news at this time. Money, power, and political influence motivate different groups to create computer-driven means of human control.

Notes

1. Andrew Zaleski, "How Bots, Twitter, and Hackers Pushed Trump to the Finish Line," Backchannel, *Wired*, November 10, 2016, https://www.wired.com/2016/11/how-bots-twitter-and-hackers-pushed-trump-to-the-finish-line/; Alessandro Bessi and Emilio Ferrara, "Social Bots Distort the 2016 U.S. Presidential Election Online Discussion," *First Monday* 21, no. 11 (November 7, 2016), http://journals.uic.edu/ojs/index.php/fm/rt/printerFriendly/7090/5653.
2. Tony Haile, "What You Think You Know about the Web Is Wrong," Time.com, March 9, 2014, http://time.com/12933/what-you-think-you-know-about-the-web-is-wrong/.
3. Don Evon, "Nope Francis," Snopes, July 24, 2016, www.snopes.com/pope-francis-donald-trump-endorsement/.
4. Marc Fisher, John Woodrow Cox, and Peter Hermann, "Pizzagate: From Rumor, to Hashtag, to Gunfire in D.C.," *Washington Post*, December 6, 2016, https://www.washingtonpost.com/local/pizzagate-from-rumor-to-hashtag-to-gunfire-in-dc/2016/12/06/4c7def50-bbd4-11e6-94ac-3d324840106c_story.html.
5. D. Parvaz, "The Arab Spring, Chronicled Tweet by Tweet," Al Jazeera English, November 6, 2011, www.aljazeera.com/indepth/features/2011/11/2011113123416203161.html; Sara El-Khalili, "Social Media as a Government Propaganda Tool in Post-revolutionary Egypt," *First Monday* 18, no. 3 (March 4, 2013), http://firstmonday.org/ojs/index.php/fm/rt/printerFriendly/4620/3423.
6. "Twitter Served as a Lifeline of Information During Hurricane Sandy," Pew Research Center, FactTank, October 28, 2013, www.pewresearch.org/fact-tank/2013/10/28/twitter-served-as-a-lifeline-of-information-during-hurricane-sandy/.
7. David Turner, Michael Schroeck, and Rebecca Shockley, *Analytics: The Real-World Use of Big Data in Financial Services*, executive report (Somers, NY: IBM Global Services, 2013).
8. Kate Crawford, "The Hidden Biases in Big Data," *Harvard Business Review*, April 1, 2013, https://hbr.org/2013/04/the-hidden-biases-in-big-data.
9. Dirk Helbing, Bruno S. Frey, Gerd Gigerenzer, Ernst Hafen, Michael Hagner, Yvonne Hofstetter, Jeroen van den Hoven, Roberto V. Zicari, and Andrej Zwitter, "Will Democracy Survive Big Data and Artificial Intelligence?" *Scientific American,* February 25, 2017, https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/; previously published in *Scientific American's* sister publication *Spektrum der Wissenschaft* as "Digitale Demokratie statt Datendiktatur."
10. Steven J. Frenda, Rebecca M. Nichols, and Elizabeth F. Loftus, "Current Issues and Advances in Misinformation Research," *Current Directions in Psychological Science* 20, no. 1 (2011): 20–23.
11. Haile, "What You Think You Know."
12. Igal Zelfman, "Bot Traffic Report 2016," *Imperva Incapsula Blog*, January 24, 2017, https://www.incapsula.com/blog/bot-traffic-report-2016.html.
13. Onur Varol, Emilio Ferrara, Clayton A. Davis, Filippo Menczer, and Alessandro Falmmini, "Online Human-Bot Interactions: Detection, Estimation and Characterization," in *Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017)* (Palo Alto, CA: AAAI Press, 2017), 280.
14. Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini, "The Rise of Social Bots," *Communications of the ACM* 59, no. 7 (July 2016): 96.
15. Philip N. Howard, *Pax Technica: How the Internet of Things May Set Us Free or Lock Us Up* (New Haven, CT: Yale, 2015), 211.
16. Twitter, Inc., Form 10-Q, Report for the Quarterly Period Ended June 30, 2014, US Securities and Exchange Commission file number 001-36164, www.sec.gov/Archives/edgar/data/1418091/000156459014003474/twtr-10q_20140630.htm; Varol et al., "Online Human-Bot Interactions.
17. Charles F. Bond and Bella M. DePaulo, "Accuracy of Deception Judgments," *Personality and Social Psychology Review* 10, no. 3 (2006): 214–34.
18. Niall J. Conroy, Victoria L. Rubin, and Yimin Chen, "Automatic Deception Detection: Methods for Finding Fake News," *Proceedings of the Association for Information Science and Technology* 52, no. 1

(2015), https://doi.org/10.1002/pra2.2015.145052010082.
19. Jeffrey Hancock, Michael T. Woodworth, and Stephen Porter, "Hungry like the Wolf: A Word-Pattern Analysis of the Languages of Psychopaths," *Legal and Criminological Psychology* 18 (2013): 102–14; David M. Markowitz and Jeffrey T. Hancock, "Linguistic Traces of a Scientific Fraud: The Case of Diederick Stapel," *PLOS ONE* 9, no. 8 (2014), https://doi.org/10.1371/journal.pone.0105937; Rada Mihalcea and Carlo Strapparava, "The Lie Detector: Explorations in the Automatic Recognition of Deceptive Language" (short paper, Joint Conference of the 47th Annual Meeting of the Association for Computational Linguistics and 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, Singapore, August 2–7, 2009).
20. Momchil Hardalov, Ivan Koychev, and Preslav Nakov, "In Search of Credible News," in *Artificial Intelligence: Methodology, Systems, and Applications: 17th International Conference, AIMSA 2016, Varna, Bulgaria, September 7–10, 2016, Proceedings*, ed. C. Dichev and G. Agre (London: Springer, 2016), 172–80; Markowitz and Hancock, "Linguistic Traces," E105937; Milhalcea and Strapparava, "The Lie Detector."
21. Song Feng, Ritwik Banerjee, and Yejin Choi, "Syntactic Stylometry for Deception Detection," in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics* (New York: Association for Computational Linguistics, 2012), 171–75, www.aclweb.org/anthology/P12-2034.
22. Victoria L. Rubin and Tatiana Lukoianova, "Truth and Deception at the Rhetorical Structure Level," *Journal of the Association for Information Science and Technology* 66, no. 5 (2015): 905–17.
23. Jacob Ratkiewicz, Michael Conover, Mark Meis, Bruno Goncalves, Snehal Patil, Alessandro Flammini and Filippo Mercer, "Truthy: Mapping the Spread of Astroturf in Microblog Streams," In *WWW '11: Proceedings of the 20th International Conference Companion on World Wide Web* (New York: Association of Computational Linguistics, 2011), 249–52, http://doi.org/10.1145/1963192.1963301; Zhiwei Jin, Juan Cao, Yongdong Zhang, Hianshe Zhou, and Qi Tian, "Novel Visual and Statistical Image Features for Microblogs News Verification," *IEEE Transactions on Multimedia* 19, no. 3 (March 2017): 598–608.
24. Victorial L. Rubin, Yimin Chen, and Niall J. Conroy, "Deception Detection for News: Three Types of Fakes," in *ASIST 2015: Proceedings of the 78th ASIS&T Annual Meeting*, ed. Andrew Grove (Silver Spring, MD: Association for Information Science and Technology, 2015); Myle Ott, Claire Cardie, and Jeffrey T. Hancock, "Negative Deceptive Opinion Spam," in *The 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*: *Proceedings of the Main Conference* (Stroudsburg, PA: Association of Computational Linguistics, 2013), 497–501; Xin Luna Dong, Evgeniy Gabrilovich, Kevin Murphy, Van Dang, Wilko Horn, Camillo Lugaresi, Shaohua Sun, and Wei Zhang, "Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources," *Proceedings of the VLDB Endowment*, arXiv:1502.03519v1 [cs.DB] February 12, 2015.
25. Giovanni Luca Ciampaglia, Prashant Shiralkar, Luis M. Rocha, Johan Bollen, Fillippo Menczer, and Alessandro Flammini, "Computational Fact Checking from Knowledge Networks," *PLOS ONE* 10, no. 6 (2015), https://doi.org/10.1371/journal.pone.0128193.
26. Hamdi Yahuaoui Al-Mutairi and Hazem Raafat, "Lattice-Based Ranking for Service Trust Behaviors," *Knowledge Based Systems* 102 (2016): 20–38; Carlos Castillo, Marcelo Mendoza, and Barbara Poblete, "Predicting Information Credibility in Time-Sensitive Social Media," *Internet Research* 29, no. 5 (2013): 560–88.
27. Benjamin Paul Chamberlain, Clive Humby, and Marc Peter Deisenroth, "Probabilistic Inference of Twitter Users' Age Based on What They Follow," Association for the Advancement of Artificial Intelligence, arXiv:1601.04621v2 [cs.SI], February 24, 2017.
28. Aditi Gupta, Ponnurangam Kumaraguru, Carlos Castillo, and Patrick Meier, "TweetCred: Real-Time Credibility Assessment of Content on Twitter," in *Social Informatics: SocInfo 2014*, ed. L. M. Aiello and D. McFarland (London: Springer, 2014), 228–43.
29. Zhao Liang, Ting Hua, Chang-Tien Lu, and Ing-Ray Chen, "A Topic-Focused Trust Model for Twitter," *Computer Communications* 76 (2016): 1–11; Victoria L. Rubin, Niall J. Conroy, and Yimin Chen, " Towards News Verification: Deception Detection Methods for News Discourse" (paper, Hawaii International Conference on System Sciences [HICSS48] Symposium on Rapid Screening Technologies, Deception Detection and Credibility Assessment Symposium, Kauai, HI, January 2015), http://works.bepress.com/victoriarubin/6/; Rubin, Chen, and Conroy, "Deception Detection for News"; Diego Saez-Trumper, "Fake Tweet Buster: A Webtool to Identify Users Promoting Fake News on Twitter," in *HT '14: Proceedings of the 25th ACM Conference on Hypertext and Social Media* (New York: Association for Computing Machinery, 2014), 316–17, https://doi.org/10.1145/2631775.2631786; Chen Yimin, Niall J. Conroy, and Victoria L. Rubin, "News in an Online World: The Need for an 'Automatic Crap Detector,'" *Proceedings of the Association for Information Science and Technology* 52, no. 1 (2015), https://doi.org/10.1002/pra2.2015.145052010081.
30. Clayton A. Davis, Giovanni Luca Ciampaglia, Luca Maria Aiello, Keychul Chung, Michael D. Conover, Emilio Ferrara, Alessandro Flammini, et al., "OSoMe: The IUNI Observatory on Social Media," preprint, PeerJ Preprints, accepted April 29, 2016, https://doi.org/10.7287/peerj.preprints.2008v1.
31. V. S. Subrahmanian, Amos Azaria, Skylar Durst, Vadim Kagan, Aram Galstyan, Kristina Lerman, Linhong Zhu, et al., "The DARPA Twitter Bot Challenge," *Computer,* June 2016, 38.
32. Norah Abokhodair, Daisy Yoo, and David W. McDonald, "Dissecting a Social Botnet: Growth, Content and Influence in Twitter," *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing* (New York: Association for Computing Machinery, 2015), 839–51, https://doi

.org/10.1145/2675133.2675208; Lorenzo Alvisi, Allen Clement, Alessandro Epasto, Silvio Lattanzi, and Alessandro Panconesi, "SoK: The Evolution of Sybil Defense via Social Networks," in *Proceedings of the 2013 IEEE Symposium on Security and Privacy* (Piscataway, NJ: Institute of Electrical and Electronics Engineers, 2013), 382–96, https://doi.org/10.1109/SP.2013.33; Yazan Boshmaf, Ildar Muslukhov, Konstantin Beznosov, and Matei Ripeanu, "The Socialbot Network: When Bots Socialize for Fame and Money" (paper, 27th annual Computer Security Applications Conference, ACSAC 2011, Orlando, FL, December 5–9, 2011); Qiang Cao, Xiaowei Yang, Jieqi Yu, and Christopher Palow, "Uncovering Large Groups of Active Malicious Accounts in Online Social Networks" in *CCS '14: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security* (New York: ACM: 2014), 477–88, https://doi.org/10.1145/2660267.2660269; Clayton Allen Davis, Onur Varol, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer, "BotOrNot: A System to Evaluate Social Bots," in *WWW '16 Companion: Proceedings of the 25th International Conference Companion on World Wide Web*, 273–74, https://doi.org/10.1145/2872518.2889302; Chad Edwards, Autumn Edwards, Patric R. Spence, and Ashleigh K. Shelton, "Is That a Bot Running the Social Media Feed?" *Computers in Human Behavior* 33 (2014) 372–76; Aviad Elyashar, Michael Fire, Dima Kagan, and Yuval Elovici, "Homing Social Bots: Intrusion on a Specific Organization's Employee Using Socialbots" in *ASONAM '13: Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (New York: ACM, 2013), 1358–65, https://doi.org/10.1145/2492517.2500225; Carlos Freitas, Fabricio Benevenuto, Saptarshi Ghosh, and Adiano Veloso, "Reverse Engineering Socialbot Infiltration Strategies in Twitter." arXiv:1405.4927 [cs.SI], May 20, 2014; Russell Frank, "Caveat Lector: Fake News as Folklore," *Journal of American Folklore* 128, no. 509 (Summer 2015): 315–32; Varol et al., "Online Human-Bot Interactions"; Claudia Wagner, Silvia Mitter, Christian Körner, and Markus Strohmaier, "When Social Bots Attack: Modeling Susceptibility of Users in Online Social Networks" in *Making Sense of Microposts: Proceedings of the WWW '12 Workshop on "Making Sense of Microposts,"* ed. Matthew Row, Milan Stankovic, and Aba-Sah Dadzie (CEUR Workshop Proceedings, 2012), http://ceur-ws.org/Vol-838.

33. Claire Wardle, "Fake News: It's Complicated," First Draft News, February 16, 2017, https://medium.com/1st-draft/fake-news-its-complicated-d0f773766c79.
34. Samuel C. Woolley and Douglas R. Guilbeault, *Computational Propaganda in the United States of America: Manufacturing Consensus Online,* Computational Propaganda Research Project, Working Paper 2017.5 (Oxford, UK: Project on Computational Propaganda, 2017), 8, http://comprop.oii.ox.ac.uk/2017/06/19/computational-propaganda-in-the-united-states-of-america-manufacturing-consensus-online/.
35. Woolley and Guilbeault, *Computational Propaganda*, 22.
36. Douglas Guilbeault, "Growing Bot Security: An Ecological View of Bot Agency," *International Journal of Communication* 10 (2016): 5012.
37. Guilbeault, "Growing Bot Security," 5012.
38. Guilbeault, "Growing Bot Security," 5003–21.
39. Samuel C. Woolley and Philip N. Howard, "Political Communication, Computational Propaganda, and Autonomous Agents." *International Journal of Communication* 10 (2016): 4882–90; Sam Woolley and Phil Howard, "Bad News Bots: How Civil Society Can Combat Automated Online Propaganda," TechPresident, December 10, 2014, http://techpresident.com/news/25374/bad-news-bots-how-civil-society-can-combat-automated-online-propaganda; Jonathan Stray, "Defense against the Dark Arts: Networked Propaganda and Counter-propaganda." Jonathan Stray website, February 24, 2017, http://jonathanstray.com/networked-propaganda-and-counter-propaganda.
40. Tao Stein, Ergond Chen, and Karan Mangla, "Facebook Immune System," in Proceedings of the 4th Workshop on social network systems. Article #8. *EuroSys Social Networks Systems (SNS)* 2011, April 10, 2011 Salzburg, http://www.cse.iitd.ac.in/~siy107537/sil765/readings/a10-stein.pdf.
41. Charles Warner, "Google Increases Regulation of False Ads and Fake News," *Forbes*, January 25, 2017, https://www.forbes.com/sites/charleswarner/2017/01/25/google-increases-regulation-of-false-ads-and-fake-news/.
42. Emily Bell, "Facebook Drains the Fake News Swamp with New, Experimental Partnerships," Little Green Footballs, December 15, 2016, http://littlegreenfootballs.com/page/322423_Facebook_Drains_the_Fake_News_.
43. Kathleen Chaykowski, "Facebook Expands Fight against Fake News with Automatic, Related Articles," *Forbes,* August 3, 2017, https://www.forbes.com/sites/kathleenchaykowski/2017/08/03/facebook-expands-fight-against-fake-news-with-automatic-related-articles/.