THE SPREAD OF TOP MISINFORMATION ARTICLES ON TWITTER IN 2017: SOCIAL BOT INFLUENCE AND MISINFORMATION TRENDS

by

Alyssa Schlitzer

A Thesis Submitted to the Faculty of the

SCHOOL OF JOURNALISM

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF ARTS

In the Graduate College

THE UNIVERSITY OF ARIZONA

2018

STATEMENT BY AUTHOR

The thesis titled ***The Spread of Top Misinformation Articles on Twitter from 2017-2018: Social Bot Influence and Misinformation Trends*** prepared by ***Alyssa Schlitzer*** has been submitted in partial fulfillment of requirements for a master's degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: alyssa Schlitz

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

Jeannine Relly, Ph.D.
Associate Professor of Journalism

Defense date
5-1-2018

TABLE OF CONTENTS

APPENDICES

With Gratitude

I wish to express my earnest gratitude to the chair of my thesis committee, Professor Jeannine Relly, for her kindness, utmost respect and dedication to helping me succeed during my master's program. It was a true honor to have such an incredible researcher, leader and mentor to encourage me along the way. I would especially like to thank Professor Relly for her constant encouragement, humble attitude and thoughtful suggestions throughout my thesis. She taught me how to think critically, ask tough questions during the research process and develop better time management skills. Professor Relly was a big inspiration to me during my undergraduate career in journalism at the University of Arizona, and was even more so during my graduate career.

I also would like to thank committee members Professor Carol Schwalbe, Professor Michael McKisson and Professor Mary Feeney, who inspired me to conduct research on the ground-breaking issues related to the spread of misinformation on social media by bots. All of these professors have inspired me in different ways during my college career, and I am thankful for their dedication in helping me succeed. My thesis would not have been possible without the help from all of these talented, intelligent and inspiring professors.

Dedication

I am grateful for my incredible husband, Levi Hill, who motivated me along the way. I could not have completed my thesis without his encouragement and patience. I also would like to dedicate my thesis to my family and grandparents who helped make my dream come true. They are all my reason for not giving up, even when my thesis seemed impossible to finish.

Abstract

Misinformation and how it spreads on Twitter is important for understanding the automated realm our world has developed into over time. Scholarship has focused on manipulation of information, but less research has focused on misinformation on Twitter. This thesis analyzed how social bots influenced the spread of top misinformation articles compared to non-bot users on Twitter in 2017. The study analyzed top misinformation and fact-checking article topic trends from January 1, 2017 to December 31, 2017, in the weeks leading up to and during the first year of Donald J. Trump's presidency in the U.S. The top articles refer to those that spread the farthest on Twitter in 2017. The case study utilized descriptive statistics to analyze trends in the data and a qualitative content analysis to establish topics and subtopics of misinformation articles ($N = 480$). For misinformation article topics, referring to the most popular topics spread by users in 2017, "general politics" and "Trump" were the most common topic and sub-topic. For fact-checked articles, the most common article topic and subtopic were also "general politics" and "Trump." The most common misinformation sources were online "news" outlets Breitbart and Infowars. While this doesn't suggest that these are "fake news" websites, the findings show that these "news" outlets regularly post misleading information. Misinformation articles also spread more rapidly on Twitter than fact-checked articles did. This suggests that users spread misinformation more commonly, rather than fact-checked articles. Lastly, the "most active users" of misinformation articles on Twitter were examined to discover social bot trends. The findings show that social bots impacted the spread of misinformation articles, mostly related to topics about "Trump" and "general politics."

Keywords: misinformation, disinformation, computational propaganda, fake news, social bots, Twitters, content analysis, automated users

Introduction

In August 2017, a false story by American conspiracy theorist "news" website, Infowars, which was shared more than two million times on social media, claimed the first Category 6 Hurricane hit the Americas, even though category 6 hurricanes do not exist (Rannard, 2017). This was not the first time "fake news followed the biggest news stories of 2017" (Rannard, 2017, para. 3). This is because misinformation spreads virally on social media, where 62 percent of American adults get their news (Gottfried et al., 2016). The amount of news Americans consume from Twitter increased from 59 percent in 2016 to 74 percent the following year, and it is projected to increase more in 2018 (Gottfried et al., 2016). The spread of misinformation on Twitter is increasing with the rise of social bots that spread unreliable articles around social networks (Bessi et al., 2016). This thesis, which studies misinformation on social media, utilizes Ecker et al.'s (2014) definition of misinformation in news articles, that is, misleading information that contains at least one piece of false information. Misinformation comes in many different forms, including hoaxes, rumors, conspiracy theories, propaganda and fake news. Misleading information has become worrisome because misinformation spreads on social media more rapidly now than ever before (Ciampaglia, 2017).

Social bots, are changing the way that news is spread, understood, and conceptualized in modern day society (Zhao, 2006). In the digital age, the use of social bots has become more popular with the increase of social media use, computer technology, digital news production and the dissemination of articles on social networks. Social bots are computer algorithms that are programmed to copy human behavior on the Internet by posting, sharing and generating information through different social networking platforms, such as Facebook and Twitter. For this thesis, I examined top news articles containing misinformation between January 1, 2017, and

December 31, 2017, and how they spread on Twitter. These "top" articles are those that spread the farthest on Twitter in 2017. The second part of my research focused on how social bots affected the spread of these top misinformation articles in the year 2017. The case study included collecting a census of data and analyzing trends in misinformation articles, the types of articles bot accounts are more susceptible to spreading and which "news sources" and article topics received the most engagement on Twitter in 2017. Yin (2017) writes case studies are, "a social science research method, generally used to investigate a contemporary phenomenon in depth and its real-world context" (p. 286). Engagement refers to the amount of times a user interacts with a tweet, including retweets, replies and likes (Lu, 2014).

This research is timely because the spread of misinformation on Twitter is increasing and is changing how the public digests news. Bots also have played a vital role in the spread of fake news and misinformation as social media use increases around the world each year. The time period was chosen for this research because it is known that misinformation was spread, with the help of social bots, leading up to the 2016 presidential election and may have led Donald Trump to victory (Desigaud et al., 2017). It is not known how the misinformation article topics changed on social media during the first weeks of 2017, as President-elect Trump readied to take office or during the presidency that year from January 20, 2017 to December 31, 2017. The following sections focus on the concepts, and differences, among fake news, misinformation, computational propaganda and disinformation. Further, the thesis explores how social bots on Twitter are constructed and the impacts they have on the spread of misinformation on social media. Previous scholarly work on these concepts also is explored in the following sections.

Literature Review

Fake news, misinformation, computational propaganda, and disinformation

The rapid rise of fake news is a major issue for society. It still is unclear why websites that generate fake news are making more content now than ever (Vargo et al., 2016). Studies show that younger generations tend to be most drawn to opinionated or fake news because they are interested in more “authentic” stories and fake news delivers narratives that fit the need (Marchi, 2012). One study, for example, included 61 interviews with racially diverse high school students and found that many of them are not interested in “objective” news, because they have the “desire to gain a more balanced understanding of news that attracted them to blogs, Facebook postings, YouTube videos, fake news, and other nontraditional sources of news” (Marchi, 2012, p. 256). The Internet provides a large selection of online sources, where factual reporting is often displaced with “alternative” narratives, or more specifically “fake news,” which portrays the media watchdogs as those who deliberately misinform (Albright, 2017). For the purposes of this study, fake news will be defined as “low quality news with intentionally false information” (Shu et al., 2017, p. 1). Misinformation refers to incorrect information in news articles that may be misleading and contain at least a piece of false information (Ecker et al., 2014). Misinformation, for example, includes hoaxes, rumors, conspiracy theories, propaganda and “fake news” (Ciampaglia, 2017). According to Woolley et al. (2016), computational propaganda “is the assemblage of social media platforms, autonomous agents, and big data tasked with the manipulation of public opinion,” and, “involves software programs that are interactive and ideologically imbued.” (p. 5). Disinformation is intentionally false information with the purpose to mislead the user (Tudjman et al., 2003, p. 2). “Misinformation (inaccurate information) and disinformation (deceptive information) easily diffuse, over time, across social groups. Social media, such as Twitter and Facebook, have made dissemination and diffusion easier and faster" (Karlova et al., 2013, p. 2).

With today's elaborate infrastructure of technologies now available, social media has created an ideal environment for manipulation (Waldrop, 2017). For example, content created from fake news websites increased from almost 1,000 articles in January 2014 to over 6,000 fake news articles in 2016 (Vargo et al., 2016). Research has demonstrated that people with high exposure to fake news perceive the "realism" of fake news more and real news less than those who have high exposure to both fake and real news, especially in relation to political satire (Balmas, 2012). The spread of misinformation during political elections is a strong example of how misinformation, from non-journalistic entities and major news outlets, has affected the accountability of the media in the public's eye (Waldrop, 2017). By the 2016 U.S. presidential election, the trickle of misinformation on social media "had become a tsunami" because social spam evolved into "political clickbait," which is "fabricated money-making posts that lured millions of Facebook, Twitter, and YouTube users into sharing provocative lies" (p. 12631).

Computational propaganda, such as the use of social media to influence public perceptions, thrived during the 2016 U.S. presidential campaign because misinformation was spread online with the goal to mislead voters through the spread of fake news that may have led Donald Trump to victory (Desigaud et al., 2017). Fake news sites use social media to attract web traffic, increase user engagement and use social media as a crucial tool for digital propaganda efforts to generate false support (Desigaud et al., 2017). Disinformation, which is information that is deliberately created to mislead the public, and propaganda from partisan sites operated by both political parties played a big role in the election (Faris et al., 2017), but it was more rampant on the right than on the left, according to research, “as it took root in the dominant partisan media on the right, including Breitbart, Daily Caller, and Fox News” (p. 21).

Since Donald Trump was inaugurated as the U.S. president on January 20, 2017, he has used the term “fake news” as a rhetorical tool to intimidate opponents, and many political leaders around the world have utilized the phrase as a tool for fighting back against critics and damaging institutions of democracy (Erlanger, 2017). In threatened countries or those with a restricted press, such as Russia, Turkey, China, and Somalia, “political leaders have invoked ‘fake news’ as justification for beating back media scrutiny” (Erlanger, 2017, para. 5).

Past research examined the role of misinformation in society and on social media, as well as the motivation behind its creation. One researcher’s online experiment with randomly selected, English-speaking U.S. adults found misinformation corrections were effective, even in the presence of partisan motivations and emotional experiences (Weeks, 2015). A group of researchers also conducted a topical study to determine how misinformation affects the public’s perception of health topics, such as cancer, including misinformation articles about e-cigarettes, indoor tanning, reusing plastic bottles, and artificial sweeteners, impact the public’s perception on these issues (Tan et al., 2015). Researchers conducted a web-based survey experiment to determine if individuals’ attitudes toward misinformation on GMOs is more difficult to change than public debate issues regarding vaccines and autism and found social media might play a part in “propagating or correcting misinformation” (Bode et al., 2015, p. 634). According to Southwell et al’s work, “Some types of misinformation—and misinformation about some topics—are more consequential than others when considering the possible negative impacts for society” (2015, p. 592). Even though researchers are able to collect more data than ever before, the data that explains how fake news spreads and impacts the public has been understudied

(Albright, 2017). Further research about "the recent advances in algorithmic content creation" that has "allowed fake news to automate news stories" is the next step in determining how fake news websites are creating more content every year (Vargo et al., 2016, p. 16).

Hoaxy, a new platform created by researchers at Indiana University that tracks how articles containing false or inaccurate information spreads on social media, has already been successful in monitoring the spread of disinformation online regarding a Russian influence campaign against the White Helmets in rebel-controlled Turkey and Syria, also known as the Syria Civil Defence (Solon, 2017). By searching the keywords "White Helmets" in Hoaxy, a handful of sources were noted that generated hundreds of stories about the organization (Solon, 2017). Hoaxy identified Beeley, a British blogger, as one of the most "influential disseminators of content about the White Helmets" (Solon, 2017, para. 31). Hoaxy also has been used to "examine how social bots promoted hundreds of thousands of false and misleading articles spreading through millions of Twitter posts during and following the 2016 U.S. presidential campaign" (Shao et al., 2016, p. 2). Shao et al. (2016) found that misinformation sources all produced about "100 articles per week" and by the end of the study period, "the mean popularity of these claims was approximately 30 tweets per article per week" (p. 3). Botometer, a social bot tracking software application that determines the likelihood that a Twitter account is a bot, has been successfully used by researchers to determine the effects that social bots had on the spread of disinformation in the lead-up to the 2017 French presidential election (Ferrara, 2016) and during the 2016 U.S. presidential election (Bessi et al., 2016; Shao, et al., 2016). For the current case study, the research focused on how misinformation trends changed in the months after Trump was elected president and became president, rather than the election itself, and the impacts that social bots had on the spread of those articles on Twitter.

Types of social bots on Twitter

Social bots are a special category of robots defined as human-made autonomous entities that interact with humans in a humanlike way over many different online platforms (Zhao, 2006). These bots are defined as “human-made autonomous entities,” which means robotic (Zhao, 2006, p. 405). They “interact with humans” in a social and “humanlike way” (Zhao 2006, p. 405). According to Woolley (2016), “Social bots have direct communication with human users on social media platforms and elsewhere in comments sections of online news sites” (p. 1). They have become an important phenomenon on social media in recent years and there are many ways in which social bots influence online discourse, such as through spam hashtags, scam Twitter users and astroturfing (Abokhodair et al., 2015). Astroturfing refers to information that is pushed out on social media by software bots (Waldrop, 2017).

There are many different types of social bots found on social media. For example, one version of social bots, known as “news bots,” are automated accounts that participate in the dissemination of information on the Internet and work as distributors of news and misinformation articles on social media (Diakopoulos et al., 2015). Another type of bot, called “malicious” bots, are created to “mislead, exploit, and manipulate social media discourse with rumors, spam, malware, misinformation, slander, or even just noise,” which could damage news accountability or lead to social issues (Ferrara et al., 2017, p. 2). Spambots are “bots used to spread marketing information on a variety of online communication platforms,” and commentary bots, “social media bots used to send out jokes or comment on social issues,” are newer, and better designed, types of social bots (Bessi et al., 2016, para. 4). Cyborgs, which are a cross between a human and bot, are referred to as “bot-assisted human or human-assisted bot” (Chu et al., 2012, p. 1). Other types of bots include “broadcast bots,” which aim to disseminate information to general audiences "by providing, e.g., benign links to news, blogs or sites" and are usually managed by organizations or bloggers (Oentaryo, 2016, p. 95). "Consumption" bots are created to "aggregate content from various sources and/or provide update services (e.g., horoscope reading, weather update) for personal consumption or use" (Oentaryo, 2016, p. 95).

Characteristics and construction of social bots

Creating social bots "has become increasingly simpler," and "in some cases, no coding skills are required to setup accounts that perform simple automated activities" (Bessi et al., 2016, p. 3). Twitter bots are computer programs that use an Application-Programming Interface (API) that control Twitter accounts and usually run on Internet-connected servers (Kollanyi, 2016). Twitter has an open API policy, which provides third-party applications, including bots, access to Twitter's public data and the ability to control social media accounts (Agarwal, 2017). Twitter bots can post their own tweets, retweets, follow or unfollow users, create polls, post media or send direct messages (Kollanyi, 2016). Most bots act as "computationally enhanced conduits for human coders" and scrape the Internet for information, examine chat sites for misuse, and analyze databases for trends (Woolley et al., 2016, p. 1).

The public can create simple Twitter bots with online software tools. A common way to make a bot is through Google Scripts (Agarwal, 2017). First, the user gives data access to Google Scripts, and then the bot creator specifies a phrase and action for the bot (Agarwal, 2017). The bot finds all the tweets that match the criteria, performs actions based on the data found, and is automatically initialized to auto-run on Twitter (Agarwal, 2017). Bots are more likely to "speed up the process of discovery and dissemination of particular stories" than humans are because they turn unknown hashtags "into the next big thing" (Manjoo, 2017, para. 29). Social bots are capable of searching Twitter for hashtags, phrases, and keywords and automatically retweeting them. They can automatically reply to tweets that meet a certain criteria and automatically follow users who tweet a specific hashtag, phrase, or keyword (Bessi et al., 2016). Depending on how they are programmed, social bots can automatically follow back users that follow them first, follow users who follow a specified user, and add users tweeting about something in particular to public lists (Bessi et al., 2016). Bots can search Google for articles "according to specific criteria and post them, or link them in automatic replies to other users," as well as automatically aggregate "public sentiment on certain topics of discussion" and post tweets automatically (Bessi et al., 2016, p. 1). Social bots scrape photos on the Internet and tend to retweet messages from other accounts within seconds (Dewey, 2016). Using entropy as a measure of uncertainty, researchers have determined that the main difference between bots and humans is that "humans have complex timing behavior, i.e., high entropy, whereas bots and cyborgs are often given away by their regular or periodic timing, i.e., low entropy" (Chu et al., 2012, p. 29). When examining the text of tweets, researchers found "a high proportion of bot tweets contain spam content" (Chu et al., 2012, p. 10).

Bots' impact on spread of misinformation

One study found that news bots can have positive impacts on journalists because automated bots can write short data-based articles on behalf of journalists, while journalists can focus on in-depth and investigative reporting (Dalen, 2012). The influence that bots have on the news media has been studied through social experiments. For example, a team of researchers created its own bot that was capable of making groups on social media to investigate the relationship of trust, popularity, and influence that bots can have on social networks (Aiello et al., 2012). The results found that the bot was widely mistaken as a human user, thus showing that untrustworthy individuals can be influential through simple automated activity if the bot becomes popular enough (Aiello et al., 2012). People more aware of the presence of bots on social media have been more inclined to follow suggestions from bots (Aiello et al., 2012).

Bots have been studied through institutional theory, the idea that organizations are influenced by environmental factors that mirror an institution's ideologies (Napoli, 2014). For example, social bots have the capacity to copy human behaviors by posting content on social media and generating news articles from data, which reflects the ideologies of media institutions (Napoli, 2014). Bots have also been studied through development, such as the AfDStatBot on Wikipedia that posted to Twitter (@WikiWars) every time an article on Wikipedia was deleted, in efforts to examine how bots interact with a niche audience of users (Geiger, 2014). Bots have been studied to determine the effects they have on the accountability of journalism as they spread topical, niche, and local news (Lokot et al., 2015). The researchers examined 360 bots, and discovered that they provided "higher-order journalistic functions such as commentary, critique, or even accountability" (Lokot et al., 2015, p. 2).

Some news bots are built to investigate information that is hidden or hard to obtain, such as the Twitter bot @Treasury_io that generates information from badly formatted U.S. Treasury data reports, while other news bots replicate the functions of journalists by ensuring accountability of those in power (Lokot et al., 2015). For example, @yourrepsonguns retweets congressional tweets about firearms, and @NYTAnon mentions every anonymous source used in New York Times articles (Lokot et al., 2015). Researchers have examined ethical uses of "blockbots," which are designed to block certain users from Twitter for inappropriate behavior, and the issues they raise about online harassment (Geiger, 2016). Many scholars have concerns about how automated systems used for filtering content can function as invisible gatekeepers, similar to how content is controlled in the media (Geiger, 2016). A study has also explored how algorithms are challenging journalistic ethics in terms of objectivity, authority, and transparency (Dörr et al., 2016). A study of robot journalism revealed the practical, sociopolitical, legal, and occupational implications for news organizations (Montal et al., 2016). A one-day case study to determine the role of algorithmic transparency in the media showed how news outlets currently use algorithms in content creation, organization and dissemination on social media (Diakopoulos et al., 2016).

Little social science research has focused on the issues of bots in relation to misinformation. The increase in social media as a "news source" has led to the rise and spread of fake news through the creation of social bots on social media platforms (Conover et al., 2011). The use of social bots to manipulate political campaigns has been evident through the spread of misinformation. For example, during the 2010 U.S. midterm elections and 2016 presidential campaign, social bots were implemented on social networks to support candidates and discredit their opponents by creating thousands of tweets that were taken from fake news websites (Conover et al., 2011). A challenge of bots "is the fact that they can give the false impression that some piece of information, regardless of its accuracy, is highly popular and endorsed by many, exerting an influence against which we haven't yet developed antibodies" (Ferrara et al., 2017, p. 2). One study explored how social bots are different from other bots and discovered they can be categorized into two dimensions—imitation of human behavior and intent—by examining bots that control their own accounts and how they spread news around social media (Stieglitz et al., 2017). After the 2016 U.S. presidential campaign, researchers discovered that accounts that spread misinformation were likely to be bots, and humans are vulnerable to manipulation because they retweet bots (Shao et al., 2017). The researchers found that 100 articles were spread per week, mostly through retweets, during the campaign from 122 websites that routinely post fake news (Shao et al., 2017). The results showed that "super-spreaders of fake news are social bots that automatically post links to articles, retweet other accounts, or perform more sophisticated autonomous tasks, like following and replying to other users" (Shao et al., 2017, p. 5).

Research Questions

Based on the literature, this study explored which misinformation topics social bots are more prone to spreading, which has been understudied. this study examined the content of the top misinformation articles on Twitter, how they spread on Twitter between January 1, 2017, and December 31, 2017, during the first year of Trump's presidency, and the extent to which social bots affected the spread of these articles. This time period was chosen because the use of Twitter is increasing each year, and the spread of misinformation on Twitter also is increasing with the rise of social bots. Other scholars have examined the spread of misinformation and the impact social bots have on the spread of those articles on Twitter during presidential campaigns. This research, on the other hand, aims to discover how misinformation trends changed during the first year of Trump's presidency. This study addresses the following research questions.

- RQ1: Which misinformation sources were the most popular during the study period?
- RQ2: Which misinformation article topics were the most popular during this year?
    - RQ2A: To what extent, if at all, did the topic trends change in the study period?
    - RQ2B: What were the most popular top fact-checking article topics during the study period?
- RQ3: What is the difference in the spread, if any, of the top misinformation[1] articles compared to the top fact-checking articles?

- RQ4: To what extent, if at all, did the most active social bots[4] help spread the top misinformation articles on Twitter compared to non-bot users?
- RQ5: What are the characteristics[5] of the top Twitter bot accounts that had the highest engagement in the spread of misinformation articles during the study period?

Method

This research focused on discovering a new framework in the relationship among misinformation articles, topics of those articles, "news" sources the articles came from, and the influence social bots had on the spread of the top misinformation articles each month from January 1, 2017 to December 31, 2017. This date range was chosen because Twitter is "where political messaging and disinformation get digested, packaged and widely picked up for mass distribution to cable, Facebook and the rest of the world. This role for Twitter grew "more intense during (and since) the 2016 campaign" (Manjoo, 2017, p. 11) and thus this research looks at the period in the year following that campaign. This date range also was chosen because President Trump has used the term "fake news" widely since he became president on January 20, 2017. For example, Trump has been "bending 'fake news' into a catchall pejorative for information he finds unfavorable" (Borchers, 2017, para. 9). The goal of my research was to study the topic areas pushed out in these misinformation articles, the most active spreaders of misinformation and how the top misinformation articles spread on Twitter during the period of the study.

Overview of Hoaxy credibility

This study utilized secondary data from the social news-sharing tracker, Hoaxy, as a starting point. This web platform allows researchers, journalists, and the public to monitor the creation of online misinformation and its corresponding fact-checking article (Shao et al., 2016).

The goal of Hoaxy is to reconstruct diffusion networks, which are the mechanisms of how information propagates through complex networks, induced by hoaxes, as well as corrections during the online sharing process (Shao et al., 2016). Hoaxy visually tracks the spread of individual misinformation articles between users on Twitter (Shao et al., 2016). According to Reaney (2016), "Hoaxy does not determine whether a story is real but it shows how it is spread online and shows related fact-checking" (para. 6). Hoaxy is still in "beta" phase, the second phase of software testing, because it is a new misinformation tracking website created by computer scientists and researchers at Indiana University in 2016. Hoaxy was awarded $1 million by the Knight Foundation in 2017 to fight the spread of misinformation (Mullin, 2017). Hoaxy's fact-checking website collection is based on the Duke Reporters' Lab database, which is constantly updated, and currently shows over 70 fact-checking websites that are active around the world (Shao et al., 2016).

For "claim" [6] misinformation sources, Hoaxy researchers compiled a list of 122 "claim" sources that spread inaccurate information, misinformation and disinformation (Shao et al., 2016). In order to continually update the list of misinformation sources, Hoaxy "implemented a tracker for the Twitter API, and a set of crawlers for both fake news and fact checking websites, as well as a database" (Shao et al., 2016, p. 2). More specifically, when a new website is added to Hoaxy's list of monitored sources, the researchers "perform a 'deep' crawl[5] of its link structure using a custom Python spyder[6] written with the Scrapy[9] framework" and also "identify

[6] Hoaxy – Sources that routinely post articles containing misinformation, misleading information, inaccurate information or disinformation.
[7] Deep Crawl – full scan of a website to scrape data (Shao, et al., 2016).
[8] Python Spyder – https://pythonhosted.org/spyder/
[9] Scrapy – https://scrapy.org/
[10] RSS feed – short for "Really Simple Syndication," which is used to gather headlines from websites and feed them directly into a database.

the URL of the RSS[10] feed, if available" (Shao et al., 2016, p. 2). After the stories are acquired, the researchers perform a crawl every two hours by checking the RSS feed (Shao et al., 2016, p. 2). Hoaxy researchers "store all these structured data into a database" (Shao et al., 2016, p. 2).

The public website, found at hoaxy.iuni.iu.edu, "can be used by reporters, researchers and the public. If a user suspects a story is false they can search it in the website to see how it was spread and to what degree it went viral" (Reaney, 2016, para. 7). The Hoaxy system collects data from both news websites and social media. For the news sources, Hoaxy first gathers data about the origin and evolution of fake news stories and their corresponding fact-checking article; then it gathers the URLs of the stories that are shared online (Shao et al., 2016). To gather data from these sources, the creators of Hoaxy used "Web scraping, Web syndication, and, where available, APIs of social networking platforms" and the Twitter streaming API for "real-time tracking of news sharing" (Shao et al., 2016, p. 2). Through this data collection, Hoaxy created a database of "claim" articles, which are the articles containing misinformation, and "fact-checking" articles that correlate with the "claim" articles based on content and URL matches through Web scraping and Twitter's Streaming API (Shao et al., 2016).

Hoaxy does not perform its own fact checking. Instead it is programmed to track news articles from sources that were determined independently from reliable fact-checking websites, such as Snopes, PolitiFact, and FactCheck (Ferrara, 2017; Shao et al., 2016; Tam, 2016). Hoaxy tracks the spread of "unverified claims," which are articles that may contain misinformation, and fact-checking articles on social media. To test the accuracy of cross-correlation between "claims" and "fact-checking" articles, the academic creators of Hoaxy gather a small sample of keywords and use them to match patterns on the lists of URLs (Shao et al., 2016). The Hoaxy researchers fetch the page that is linked in the tweet and store the URL and the article text in the database (Shao et al., 2016).

Selecting the article sample with Hoaxy

Using a census of the Hoaxy's database from January 1, 2017, to December 31, 2017, data were gathered about how articles containing misinformation spread on Twitter using Hoaxy's API features. Hoaxy was used to gather the top 20 articles each month that contained misinformation and the top 20 fact-checking articles. These top 40 articles each month were the ones that received the most user engagement and spread the farthest on Twitter that month. Other articles that month most likely did not receive enough engagement on Twitter to investigate at the same level. Article topic trends and engagement differences of the top 20 "claim" misinformation articles and top 20 fact-checking articles were gathered during the study period using Hoaxy as secondary data. Additionally, 40 users were collected using Hoaxy's "TopUsers" API feature, referred to as the top spreaders of "claim" misinformation articles on Twitter, each month from January 2017 to December 2017. Next, Hoaxy's "TopArticles" feature was used to gather article data, including article names, the websites they came from, and the dates they were published. Only the top 40 articles were evaluated, including 20 "claim" misinformation articles and 20 "fact-checking" articles, each month, because that is the limit of the Hoaxy API feature. Hoaxy's data collection only dates back to the end of 2016, when Hoaxy researchers began storing data from Twitter's Streaming API and collecting a list of claim and fact-checking sources (Shao et al., 2016).

This study clubbed together the top articles and top most active users to determine to what extent the top most active users per month, also referred to as "spreaders," were involved in the spread of top misinformation and fact-checking articles each month and how many of these spreaders were bot accounts. The collection of articles for this study added up to 480 articles total for the entire year from January 1, 2017, to December 31, 2017, including “claim” and “fact-checking” articles. A database was created using SPSS to organize the top 240 claimed articles from “news” sources that post misinformation, including satire, click bait, and unreliable news, as well as 240 articles flagged as misinformation from fact-checking websites. The Hoaxy Network API feature was used to evaluate the entire social network of each individual misinformation article, including users an article spread to on Twitter. Next, the data were coded using SPSS to determine if any of the “most active” spreaders during the study period were involved in the spread of the top 240 “claim” misinformation articles and to what extent. (See Appendix A.) A limitation of Hoaxy is that it describes misinformation articles as “claims” because the researchers have stated the group cannot prove that this volume of articles contains inaccurate or unverified information. Rather, Hoaxy relies on outside fact-checking organizations, and it is up to researchers to evaluate the evidence about a claim (Shao et al., 2016).

Overview of Botometer credibility

Botometer, a social bot detection tool developed by the same creators of Hoaxy at Indiana University, is 95 percent accurate in bot detection on Twitter, with a margin of error of plus or minus five percentage points (Davis et al., 2016). Botometer is a notable project that “aims to reveal how this information is generated and broadcasted, how it becomes viral, its overall reach and how it competes with accurate information for placement on user feeds” (Bilton, 2017, para. 15). Botometer uncovers “attempts to use Internet bots to boost the spread of misinformation and shape public opinion” (Bilton, 2017, para. 15).

Selecting the Twitter user sample with Botometer

The top 40 users who spread "claim" misinformation articles were evaluated to determine if they were potential bots compared to non-bot users using the Botometer tool. Botometer was employed to identify the bot scores of every user each month to determine which users were potential bots and non-bot users. To determine potential bot spreaders, Botometer was used to track which users possess characteristics of social bots. Botometer uncovers potential bot accounts on social media by extracting "more than a thousand features in six different classes: users and friends meta-data, tweet content and sentiment, network patterns, and activity time series" of user accounts (Varol et al., 2017, p. 9). The Botometer system "is based on a supervised machine learning algorithm leveraging more than a thousand features extracted from public data and meta-data about Twitter accounts" (Shao et al., 2017, p. 14). Features in detecting bots include user metadata and friend statistics (p. 14). Temporal patterns of activity, part-of-speech and sentiment analysis, which is a user's attitude toward a particular topic, are also included in detecting bots. "The classifier is trained using publicly available datasets of tens of thousands of Twitter users that include both humans and bots of varying sophistication" (Shao et al., 2017, p. 14). No methods are currently known to unequivocally establish whether a user is a bot or human, but there are techniques to determine if a user has human-like or bot-like behaviors through a feature-based social detection that Botometer uses (Ferrara, 2017). Bots typically have stock images as profile pictures because "it requires human efforts to customize a profile, therefore bots are more likely to exhibit the default setting" (Ferrara, 2017, p. 5). Bots also have the "absence of geographical metadata" because humans typically use phones to access Twitter, which records location (Ferrara 2017, p. 5). Bots tend to post an excessive amount of tweets, with a high proportion of retweets over of original tweets (Ferrara 2017, p. 5). Lastly, bots typically have fewer followers and more "followees" (Ferrara 2017, p. 5). Ninety-two percent of accounts "that Twitter suspends for spamming activities are suspended within three days of the first post. Therefore, if a spamming bot survives one week, it is very likely to survive a long time" (Chavoshi et al., 2016, p. 19).

If a user has a bot score around 0.70, it is a clear indicator that the account may have very clear bot characteristics (Ferrara et al., 2016; Davis et al., 2016). Botometer only produces data from a Twitter account's last 200 tweets, so the bot scores fluctuate over time (Bessi et al., 2016). This study used Botometer to determine which of the top 480 most active users in 2017 were bots that spread misinformation articles each month. If some of the potential bots were active users for more than one month, the average score was chosen. For this study, if the user remained above 0.70, it was considered a bot. This score validity is based on a conservative choice to lower false negatives and false positives, which means it is less likely to mistake an account for a bot or a bot for a human (Shao et al., 2017; Varol et al., 2017; Bessi et al., 2017). Then, each potential bot was "profiled" to determine accurate bot characteristics. Each user was manually monitored on Twitter based on characteristics listed in the literature review. According to research, "some bots specifically aim to achieve greater influence by gathering new followers and expanding their social circles" (Ferrara et al., 2017, p. 4). Influential social bots "search the social network for popular and influential people and follow them or capture their attention by sending them inquiries, in the hope to be noticed" (Ferrara et al., 2017, p. 4). Social bots are capable of infiltrating popular discussions, generating topically appropriate content and automatically producing responses through natural language algorithms based on keywords that are relevant to them (Ferrara, et al., 2017).

With the data from Hoaxy and Botometer the most popular article topics were discovered through quantitative evaluation. Sites that received the most engagement and sites that published the most articles about each topic were also determined. Lastly, connections among social bots, non-bot users, article topics and misinformation sources were discovered in the data evaluation.

Coding the data

For this study, SPSS was used to code the article data (See Appendix A) and user data (See Appendix B), which were organized by variable. The researcher was the only coder throughout the duration of the study period. Articles were categorized topically for the entire year to determine to what extent social bots were involved in the spread of misinformation articles. This research took a qualitative approach to identify article topics, since article trends constantly change over time, and topics were not predetermined. Topics were determined based on keywords in headlines, content within the articles and which section the article was hosted on the publisher's website. For example, "data collection focused on the publication, the date, the news sources in the report and the topic of the report" (Altheide et al., 2013). Specific headline keywords, such as Donald Trump, terrorism and political crime, including the words "murder" or "scandal," were used to determine some categories during the study period in 2017. Other variables in the codebook included the news outlet the article originated from with a URL, the date the article was posted on Twitter, and if the article was either a "claim" containing misinformation or a "fact-checking" article. Misinformation articles were coded as 0 and fact-checking articles were coded as 1. If the article received less than 30,000 in engagement during the month monitored it was coded as 0, while articles that received over 30,000 in engagement were coded as 1 in SPSS. Each article also contained data regarding the number of tweets it had in the month it was monitored. For example, from the Hoaxy data collection, the top claim article between January 1, 2017, and February 1, 2017 was “CBS Confirms Trump Brought His Own Studio Audience To Clap For Him At CIA Speech” by Politicus USA. It received a total of 32,370 tweets in that time period. All nominal data were recoded for analysis in SPSS for basic descriptive statistics (See Appendix A.) For the most active users portion of the research, each Twitter account’s user handle, the number of tweets they had during the month they were monitored and their Twitter URL were collected. The Botometer API was connected to the Hoaxy “TopUsers” API to collect bot scores for each user during the month monitored. In the SPSS codebook, users with an average bot score between 0-0.49 were coded as 1, users with a bot score between 0.50-0.69 were coded as 2 and users with a bot score between 0.70-0.99 were coded as 3.

Defining tone in misinformation articles

In this research, tone of an article refers to whether the headline and content within the article are in favor of, against or neutral about a topic or individual. Articles with a positive tone were coded as 1, articles with a negative tone were coded as 2 and articles with a neutral tone were coded as 3. Positive articles contain favorable language, such as strong support and excitement about a political candidate, event or individual. “The more positive the tone toward a party in the news, the more voters will vote for that party – again” (Hopmann et al., p. 391). For example, a positive tone was selected if the misinformation article had positive evaluations about a political figure, such as president Donald Trump, in an attempt to make the public believe, or keep believing, positive claims about him.

A negative tone included articles with a biased headline, or misleading content, highlighting disapproval of a particular topic or individual. For example, articles about Trump were identified as negative if the content of the article was in opposition of him or included misleading information about his actions as president. Negative tone included articles aiming to weaken or belittle a particular idea, such as president Trump, terrorism and crime. Neutral was selected if the articles did not fall under these categories, and were coded as (3) neutral.

Findings

Twenty claim articles and 20 fact-checking articles were examined each month from January 1, 2017, to December 31, 2017. The database had a total of 240 "claim" misinformation articles from Hoaxy and the top 240 "fact-checking" articles from leading fact-checking organizations, such as Snopes, PolitiFact, and FactCheck, for the entire year. Of the 240 claim articles, 4.2 percent of them received over 30,000 tweets, replies, mentions, or quotes (engagement) on Twitter, and 95.8 percent received fewer than 30,000. Of the 240 claim articles that received less than 30,000 tweets, 10 percent had between 15,000 and 30,000 total tweets, and 90 percent received less than 15,000. Of the 240 fact-checking articles, all of them received fewer than 30,000 in engagement, which shows the top "misinformation" articles spread further on Twitter during the time period studied. Of the 480 spreaders, or most active Twitter users that spread misinformation articles during the study period, 104 individual users were discovered because most users appeared more than once between January 2017 and December 2017. Of the 104 top users in this study, 87 were defined as "non-bots" and 17 were defined as "bots."

Most popular misinformation sources

The first research question asked which misinformation sources were the most popular during the study period. Out of 240 "claim" misinformation articles examined in the study, the five most popular misinformation sources were Breitbart (n = 82, or 34.2%), Infowars (n = 46, or 19.2%), The Onion (n = 37, or 15.4%), Politicus USA (n = 26, or 10.8%) and Your News Wire (n = 15, or 6.3%). Other common misinformation sources included The Blaze (n = 10, or 4.2%), Twitchy (n = 9, or 3.8%), WND (n = 5, or 2.1%), Occupy Democrats (n = 3, or 1.3%), Big Partisan Report (n = 3, or 1.3%), Before Its News (n = 2, or .8%), ANO News (n = 1, or .4%) and US Uncut (n = 1, or .4%). (See Table 1.)

Table 1: Total amount of Twitter engagement for misinformation articles published from the top misinformation sources in 2017.
(*N* = 240)

Misinformation Source	# Total articles below 30,000 engagement	# Total Articles above 30,000 engagement
ANO News	1 (100%)	0
Before It's News	2 (100%)	0
Big Partisan Report	3 (100%)	0
Breitbart	80 (97.5%)	2 (2.5%)
Infowars	46 (100%)	0
Occupy Democrats	3 (100%)	0
Politicus USA	25 (96.2%)	1 (3.8%)
The Blaze	10 (100%)	0
The Onion	31 (83.8%)	6 (16.2%)
Twitchy	9 (100%)	0
US Uncut	1 (100%)	0
WND	5 (100%)	0
Your News Wire	14 (93.3%)	1 (6.7%)
Total 240	**230**	**10**

Source: Hoaxy API

Of the 10 articles that were above 30,000 in engagement on Twitter (See Table 1), The Onion published the top two articles in 2017 (16.2% of 240 articles). The first article titled, "Popular New Amazon Service Just Comes To Your House And Kills You," received 134,659 in engagement on October 2017. The second article with the highest engagement in 2017 was "'No Way To Prevent This,' Says Only Nation Where This Regularly Happens'" by The Onion, and received 88,434 in engagement on Twitter in September 2017. These two article topics are "satire" because The Onion is a satirical news website. The reason The Onion was included as a misinformation source in this study is because stories from The Onion can appear "real" to many social media users, (Griffin, 2014). Your News Wire published the third article in July 2017 with the highest Twitter engagement of 45,834 titled, "Haiti Official, Who Exposed The Clinton Foundation, Found Dead In Miami." The other seven articles with above 30,000 in engagement included four from

The Onion (satire), two from Breitbart (Trump and general politics) and one from Politicus USA (Trump). The Trump article by Politicus USA (n = 32,370 tweets), "CBS Confirms Trump Brought His Own Studio Audience To Clap For Him At CIA Speech," received over 2,00 more in engagement on Twitter than the Trump article published by Breitbart (n = 30,028) titled, "Mark Levin to GOP: Investigate Obama's 'Silent Coup' vs. Trump,"' during the study period in 2017.

The second research question asked which misinformation article topics were the most popular during the time period studied from January 1, 2017 to December 31, 2017 and which sources published the most articles about these topics. Of the 240 articles, the most popular category was general politics (n = 115, or 47.9%), which includes information about President Donald Trump, geopolitics, Hillary Clinton, The Obama Administration, the U.S. military, immigration and Russia (See Table 2.)

Table 2: Most common news outlets and tone for each article topic in "general politics" category (*N* = 115)

Topic *Article topic*	**# Articles** *(x of 115)*	**Most Common Misinformation Source**	**# Articles Published**	**Most Common Misinformation Source Tone**	**# Articles With Most Common Misinformation Source Tone**
Trump	50 (43.5%)	Breitbart	22 (44%)	Positive	18 (82%)
Obama	6 (5.2%)	Politicus USA	2 (33%)	Positive	2 (100%)
Military	2 (1.7%)	Breitbart	1 (50%)	Negative	1 (100%)
		The Onion	1 (50%)	Negative	1 (100%)
Immigration	11 (9.6%)	Infowars	5 (45.5%)	Negative	5 (100%)
Russia	12 (10.4%)	Breitbart	5 (41.7%)	Negative	5 (100%)
Politics: Geopolitics, Hillary Clinton, other	34 (29.6%)	Breitbart	19 (56%)	Negative	19 (100%)

Source: Hoaxy API

Most popular misinformation article topics

Out of the 240 articles, 20.8 percent included stories about Trump (n = 50). The top three sources that published articles about Donald Trump were Breitbart (n = 22, or 44%), Politicus USA (n = 15, or 30%) and Infowars (n = 7, or 14%). Of the 50 claim articles about Trump, 46 percent were negative (n = 23), 16 percent were neutral (n = 8) and 38% were positive (n = 19). Of the 22 articles about Trump from Breitbart, 82 percent were positive and 18 percent were neutral. These articles included headlines such as, "The Media Covered Trump -- Without Listening to Him or His Voters," which shows favor for Trump because the author was defending the president. Another example of a positive headline is, "Major Impact: President Donald Trump's First 50 Days in Office." The reporter highlighted positive impacts Trump has made during his presidency and used an opinionated voice in favor of Trump. For example, the lede of the article says, "Trump has signed 16 executive orders, met with an array of foreign leaders, delivered an unforgettable speech to congress, and made a meaningful impact on restoring the nation to prosperity" (Kew 2016, para. 1). Other positive articles from Breitbart about Trump are, "Orrin Hatch: Trump Has Been 'One of the Best Presidents I've Served Under,'" and "Unemployment Falls to a 28-Year Low in Age of Trump." These articles have a positive tone because they show favoritism, use supportive language toward Trump and make positive claims about his first year of presidency.

Neutral articles from Breitbart about Trump include, "Pelosi: Trump Dishonored God By Walking Away From Paris Accord," and, "15 Times Celebrities Envisioned Violence Against Trump and the GOP." The first article is neutral because the coverage in the article is based off what Minority Leader Rep. Nancy Pelosi said during a public speech about Trump, rather than original reporting from Breitbart. Both of these articles suggest a neutral tone because they do not directly show support for Trump but also do not imply they are against him. Of the 15 articles about Trump from Politicus USA, a left wing biased "news" website that regularly posts misleading articles, 100 percent had a negative tone. A few examples of articles from Politicus USA that contain a negative tone are, "Here's The Proof That Trump Is Mentally Unfit And Should Be Removed From The White House;" "Coal Miners Crushed As White House Admits Trump Lied About Bringing Back Coal Jobs;" "Political Suicide: Every Republican Who Votes For Trumpcare Better Start Looking For Work." These articles from Politicus USA exemplify disapproval of Trump as well as exaggerated information about him, such as "Trump is mentally unfit" to be president (See Table 3.)

Table 3: Tone from top misinformation sources that published stories about Trump ($N = 50$)

Misinformation Source	# Articles about Trump	% Articles with Negative Tone	% Articles with Neutral Tone	% Articles with Positive Tone
Breitbart	22	(0%)	(16%)	(82%)
Infowars	7	(28.5%)	(28.5%)	(42.8%)
Politicus USA	15	(100%)	(0%)	(0%)
Occupy Democrats	2	(100%)	(0%)	(0%)
Twitchy	1	(100%)	(0%)	(0%)
Your News Wire	1	(100%)	(0%)	(0%)
Big Partisan Report	2	(100%)	(0%)	(0%)
Total	50			

Source: Hoaxy API

Of the 240 "claim" misinformation articles, the second most popular article topic in 2017 was political crime (n = 36, or 15%), which includes criminal political scandals, murder and fraud. Articles related to crime were negative (n = 34, or 94.44%) and neutral (n = 2, or 5.56%). The top three sources that published articles about crime were Breitbart (n = 10, or 27.78%), Infowars (n = 12 or 33.33%), and Your News Wire (n = 5, or 13.89%). An example of a "claim" Source: Hoaxy API misinformation article with a negative tone about political crime is, "Trump Son In Law Jared Kushner Was Just Told To Lawyer Up Because He Committed A Crime," by Politicus USA (n = 6995 tweets). This article is considered "political crime" because the word is included in the headline and the article is "negative" because the author highlights Jared Kushner as a criminal. Another example of a political crime article is, "Bombshell: Donna Brazile warned off private eye on Seth Rich murder," by WND (n = 20998 tweets). This article is considered "political crime" because the word "murder" is included in the headline. The tone of the article is "negative" because the author highlights and accuses Brazile, in an opposing manner, of telling an investigator to stop investigating Seth Rich's death. Of the 240 "claim" misinformation articles, other common topics include satirical articles (n = 32 or 12.92%). The most common misinformation source that published "satire" stories was The Onion (n = 30). Of the 30 "claim" misinformation articles the Onion published that were "satire," 76.7% had a negative tone (n = 23), 16.7 percent had a neutral tone (n = 5) and 6.7 percent had a positive tone (n = 2). Of the 240 "claim" misinformation articles, other popular article topics were terrorism (n = 11, or 4.58%) and the media (n = 11, or 4.58%). An example of a misinformation article about "terrorism" is, "Pope Francis: 'Muslim Terrorism Does Not Exist*,*" by Breitbart and, "Proof: Muslims Celebrated Terror Attack in London," by Infowars. Many of the articles related to "the media" were negative and exposed news outlets and reporters in subjective ways. An example of a negative article about "the media" is, "CNN Host Reza Aslan Calls Trump 'Piece of Sh*t' for Correctly Identifying London Terror Attack," by Breitbart. Entertainment (n = 9, 3.75%), race (n = 10, or 4.17%), women's rights (n = 4, or 1.67%) conspiracy theories (n = 2, or 0.83%) and other (n = 9, or 3.75%) were other misinformation topics discovered in this study (See Table 4.)

Table 4: Most common tone and source for other popular misinformation topics
(*N* = 115)

Topic Article topic	# Articles (x of 115)	Most Common Misinformation Source	# Articles Published	Most Common Misinformation Source Tone	# Articles With Most Common Misinformation Source Tone
Crime	36	Infowars	12 (33.3%)	Negative	10 (83.3%)
Satire	31	The Onion	30 (96.8%)	Negative	23 (76.7%)
Terrorism	11	Breitbart	5 (45.5%)	Negative	4 (80%)
		Infowars	5 (45.5%)	Negative	4 (80%)
The Media	11	The Blaze	3 (27.3%)	Negative	3 (100%)
Other	9	Breitbart	6 (66.7%)	Negative	4 (66.7%)
Race	10	Infowars	6 (60%)	Negative	6 (100%)
Women's Rights	4	Infowars	2 (50%)	Positive	2 (100%)
Conspiracy	2	Before It's News	2 (50%)	Negative	2 (100%)
Total	115				

Source: Hoaxy API

In addressing RQ2X, I asked how the misinformation article topics changed during the study period. Many of the misinformation article topics remained constant throughout the entire year, and topics only fluctuated by two or three articles each month. The only article topic that fluctuated more throughout the year was articles in the "Trump" category. For example, many articles in the "Trump" category appeared in the top 20 "claim" misinformation articles in January (n = 6), February (n = 6), March (n = 5), April (n = 5), May (n = 6) and June (n = 6) 2017, when Trump was newly president. After June, less than four misinformation articles about Trump appeared in the top misinformation articles for each month. In September, there was only one article about Trump that appeared in the top articles and two in October.

Misinformation versus fact-checking article trends and engagement

Out of the 240 fact-checking articles examined during the study period, 35.8 percent were about president Donald Trump (n = 86). Further, 32.5 percent were in the "general politics" category (n = 78) and included information about Obama, immigration, the U.S. military and other. Lastly, 5.8 percent of the articles were related to the topic "political crime," (n = 14) of the 240 total fact-checking articles examined. This shows the three most popular topics among the top fact-checking articles were also the most popular for "claim" misinformation articles during the study period. From January 1, 2017 to December 31, 2017, PolitiFact (n = 134) published 55.8 percent of the top 240 most popular fact-checking articles, 33.8 percent were published by Snopes (n = 81), 7.5 percent were from Open Secrets (n = 18), 2 percent were from FactCheck (n = 5), 0.4 percent was from Climate Feedback (n = 1) and 0.4 percent was from Open Source (n = 1). The major difference between the spread of misinformation articles and fact-checking articles is that misinformation articles received more engagement throughout the year. This shows that claim articles spread more rapidly on Twitter than fact-checking articles did from January 1, 2017 to December 31, 2017, even though the top three article topics were similar. In this current study, more users on Twitter were prone to spread misinformation rather than fact-checked articles (See Table 5.)

Table 5: Total amount of engagement from top three misinformation and fact-checking article topics
(*N* = 240)

Article Type	Number of Articles Below 1,000 Tweets	Number of Articles With 1,001-15,000 Tweets	Number of Articles with 15,001+ Tweets
Fact-checked	158 (65.8%)	81 (33.8%)	1 (.4%)
Misinformation	0 (0%)	207 (86.3%)	33 (13.8%)

Source: Hoaxy API

Top spreaders of misinformation: presence of social bots, likely-bots and non-bots

The fourth research question asked to what extent, if at all, did the most active social bots help spread misinformation articles on Twitter compared to non-bot users. During the study period, 40 users involved in the spread of "claim" misinformation articles were examined each month from January 1, 2017, to December 31, 2017. It was discovered that not every user or bot was involved in spreading the top 240 claim articles, because they may have spread articles outside of the top 20 articles each month instead. In this study, users with bot scores between 0.49 or lower were determined as non-bot users, users with a bot score between 0.50-0.69 are likely-not bots, while users with a score 0.70 or above are likely-bots. Only users with a bot score of 0.70 or higher are considered "bots" in this study because it suggests that an account may have very clear bot characteristics (Ferrara et al., 2016; Davis et al., 2016).

Of the 480 spreaders examined during the study period, 104 individual users were discovered because most users appeared more than once in the months between January 1, 2017 and December 31, 2017. Of the 104 users, 16.3 percent were bots (n = 17), with a bot score of 0.70 or higher, 51 percent were likely-not bots (n = 53), with a bot score between 0.50-0.69, and

32.7% were non-bot users (n = 34) with a bot score of 0.49 or below. Out of the 104 users found in the study period, 68 users appeared 2 or more times during the 12-month study period. This means the 68 users were top "spreaders" of claim misinformation articles for two or more months between January 1, 2017, and December 31, 2017. Of the 104 users, 32.7 percent were top spreaders (n = 34) for only one month during the study period, 44.2 percent were top spreaders (n = 46) for two to nine months and 23.1 percent were top spreaders (n = 24) for 10 or more months. Of the 24 users that were spreaders for 10 or more months, 29.2 percent were likely-bot accounts (n = 7) and 45.8 percent were users (n = 11) with a bot score between 0.00 and 0.69. Lastly, of the 24 users, 25 percent were not bots (n = 6) because they had a bot score of .49 or lower. *JoeFreedomLove*, which is a bot with a bot score of .70 or higher, appeared as a spreader of "claim" misinformation articles for 11 of the 12 months during the study period (See Table 6.)

Table 6: Users and bots (bold) that were top spreaders of misinformation articles 10 or more times during the study period

User Handle (Only spreaders for 10 or more months)	Bot Score	# Times appeared as a top spreader (12 month period)
@JoeFreedomLove	**.70 +**	**n=12**
@AllAmericanGirl	**.70 +**	**n=9**
@Qkode	**.70 +**	**n=11**
@TwitchyTeam	**.70 +**	**n=8**
@Syqau	**.70 +**	**n=12**
@Yceek	**.70 +**	**n=10**
@TroyCoby	**.70 +**	**n=10**
@63red	.50 +	n=12
@BreitbartLondon	.50 +	n=12
@Politicususa	.50 +	n=12
@ClickHole	.50 +	n=12
@TheOnion	.50 +	n=12
@Mitchellvii	.50 +	n=12
@altnewsheadline	.50 +	n=11
@Nuiotwo	.50 +	n=11
@Npnikk	.50 +	n=10
@RealMuckmaker	.50 +	n=10
@Vnuek	.50 +	n=10
@Infowars	- .49	n=10
@Joej2020usa	- .49	n=11
@PrisonPlanet	- .49	n=12
@RealAlexJones	- .49	n=12
@RealJamesWoods	- .49	n=10
@StefanMolyneux	- .49	n=10

Source: Botometer

Spread of top misinformation articles: social bot influence and article trends

To determine how social bots impacted the spread of the top "claim" misinformation articles from January 1, 2017, to December 31, 2017, the Hoaxy Network API was used to gather social networking data about each article. The data includes every name of the user who posted an individual "claim" misinformation article and the amount of engagement the user received after posting the article on Twitter. Hoaxy's Network API helped reveal trends in how bot users and non-bot users interacted with the top "claim" misinformation articles similarly and differently during the study period. Although, due to the Hoaxy Network API limit, 10 articles, in April 2017, from the total 240 "claim" articles could not be retrieved, which means 230 "claim" misinformation articles were examined for this part of the study. These trends include: the top misinformation article topics the bots spread and the misinformation sources that were most commonly posted on Twitter by different bots.

Of the 104 users, 17 users were bots with a bot score of .70 or higher. Of the 17 bot users, 52.9 percent (n = 9) of them were involved in the spread of the top 230 "claim" misinformation articles. Of the 17 bot users, 47.1 percent (n = 8) of them were not involved in the spread of the top 230 "claim" misinformation articles. Rather, they were involved in the spread of "claim" misinformation articles in Hoaxy's database outside of this study's 240 "claim" misinformation article sample size. Of the nine bots involved in the spread of the top "claim" misinformation articles from January 1, 2017, to December 31, 2017, a total of 139 articles were spread. Of those nine bot users, *@JoeFreedomLove* spread 61 claim articles, *@AIIAmericanGirI* spread 24 claim articles, *@TroyCoby* spread four articles, *@TwitchyTeam* spread eight articles, *@yceek* spread eight articles, *@_Alt_Right_* spread three articles and *@hrkbenowen* spread 16 articles, *@_Breitbot_* spread nine articles and *@qkode* spread 6 articles. Of the nine bots, the most common article topic spread by 77.9 percent of bots (n = 7) was about “Trump.” Of the nine bots, the most common article topic spread by 22.2 percent of bots (n = 2) was “general politics” and 1.1 percent of bots (n = 1) spread “entertainment” articles containing misinformation (See Table 7.)

Table 7: Most common and least common article topics spread by bot users (n = 9)

User Handle	# Claim Articles Spread	Most Common Article Topic Spread	# Articles Spread Of Most Common Topic	Least Common Article Topic Spread	# Articles Spread Of Least Common Topic
@JoeFreedomLove	61	General Politics	33 (54.1%)	Race	1 (1.6%)
@_Breitbot_	9	Trump	7 (77.8)	Terrorism	1 (11.1%)
@AllAmericanGirI	24	Trump	8 (33.3%)	Terrorism	1 (4.2%)
@TroyCoby	4	Trump	2 (50%)	Satire	1 (25%)
				Race	1 (25%)
@TwitchyTeam	8	Entertainment	5 (62.5%)	Crime	1 (13%)
@yceek	8	Trump	3 (37.5%)	Crime	1 (13%)
@Alt_Right_	3	Trump	2 (66.7%)	Terrorism	1 (33.3%)
@hrkbenowen	16	Trump	6 (37.5%)	Entertainment	1 (6%)
@qkode	6	General Politics	2 (33.3%)	Satire	1 (17%)
		Trump	2 (33.3%)		

Source: Botometer

Of bots (n = 9) involved in the spread of 130 “claim” misinformation articles, 66.8 percent of bots (n = 7) most commonly spread articles from Breitbart (n = 81). Of articles spread from Breitbart (n = 81), the top bots that spread top “claim” misinformation articles from Breitbart were @JoeFreedomLove (n = 39), @AIIAmericanGirI (n = 13), @TroyCoby (n = 4), @Alt_Right_ (n = 3), @hrkbenowen (n = 16) and qkode (n = 3). Of the nine bot users involved in the spread of “claim” misinformation articles, @yceek spread articles most commonly from Infowars (n = 6) and @TwitchyTeam spread articles most commonly from Twitchy (n = 8). (See Table 8.

Table 8: Most common and least common misinformation source bots (n = 9) spread

User Handle	# Claim Articles Spread	Most Common Misinformation Source	# Articles Spread From Most Common Misinformation Source	Least Common Misinformation Source Spread	# Articles Spread From Least Common Misinformation Source
@JoeFreedomLove	61	Breitbart	39 (63.9%)	WND	3 (.5%)
@_Breitbot_	9	Breitbart	9 (100%)		
@AIIAmericanGirI	24	Breitbart	13 (54.2%)	Your News Wire	1 (4.2%)
@TroyCoby	4	Breitbart	3 (75%)	The Onion	1 (25%)
@TwitchyTeam	8	Twitchy	8 (100%)		
@yceek	8	Infowars	6 (75%)	Breitbart	2 (25%)
@Alt_Right_	3	Breitbart	3 (100%)		
@hrkbenowen	16	Breitbart	11 (68.8%)	Twitchy	1 (6%)

Source: Botometer

Top Twitter bot characteristics

Using Hoaxy's Botometer and examining the nine bot users on Twitter, characteristics of each individual bot user (n = 9) were discovered. These bots are considered "news bots" in this study because they are automated accounts that participate in the dissemination of information on the Internet and work as distributors of news and misinformation articles on social media (Diakopoulos et al., 2015).

@JoeFreedomLove is considered a "news bot" in this study because it possesses a bot score of 0.70 or higher and mainly publishes links to articles through original tweets. The user only has a 5 percent retweet ratio, which means the user rarely publishes content on Twitter through retweets, according to Hoaxy's Botometer. Based on the user's Twitter account, it is a Trump supporter, fighter of socialism and America enthusiast. It also uses a scraped photo from the Internet, which implies the user may be a bot (Ferrara, 2017). The user joined Twitter in 2014 and has a total of 574,871 tweets. *@JoeFreedomLove* follows 46,122 users and has 50,140 followers. A majority of this bot's tweets consist of links to original articles directing to the publisher's website and short descriptions in the tweet of what the article is about. According to

Hoaxy's Botometer API, *@JoeFreedomLove* posts 7,400 tweets per week on average. Many of these characteristics imply the user is a bot because *@JoeFreedomLove* exhibits "incessant activity and excessive amounts of tweets," and produces a high amount of "spam" tweets, such as links to links to misinformation articles, with low engagement (Ferrara, 2017, p. 5).

@_Breitbot_ is considered a "news bot" in this study because it possesses a bot score of .70 or higher and has a 100 percent retweet ratio, which means the user always publishes content on Twitter through retweets, according to Hoaxy's Botometer. Based on the user's Twitter account, it has a total of 51,600 tweets, follow 4,942 users and have 1,536 followers. The bot user joined Twitter in 2016 and produces about 17 tweets per week. It is unclear whether the user has a scraped photo from the Internet. Many of these characteristics imply the user is a bot because many bots have less followers and follow more users, tweet rapidly during the week and tend to retweet other users rather than post original content (Ferrara, 2017).

@AllAmericanGirl is considered a "news bot" in this study because it possesses a bot score of 0.70 or higher. The user has a 0.0 percent retweet ratio and mainly posts original tweets with links to articles outside of Twitter. Based on its Twitter account biography, the user has a scraped photo from the internet, is a conservative and states that they are not a "Russian bot," which implies this account may be a different type of bot. *@AllAmericanGirl* joined Twitter in 2013 and has 566,929 tweets, 118,142 followers and follows 76,644 users. The bot also tweets 4,200 times per week, on average, according to Hoaxy's Botometer API. Many of these characteristics imply the user is a bot because the user tweets rapidly per week, posts "spam" content, such as links to outside articles, and uses a scraped photo from the Internet (Ferrara, 2017).

@TroyCoby is considered a "news bot" in this study because it possesses a bot score of .70 or higher. The user has an 11 percent retweet ratio and mainly posts original tweets with links to articles outside of Twitter. Based on its Twitter account biography, the user has a scraped photo from the Internet, joined Twitter in 2016 and has 348,092 tweets, 19,579 followers and follows 16,465 users. The bot also tweets 6,900 times per week, on average, according to Hoaxy's Botometer API. Many of these characteristics imply the user is a bot because the user tweets rapidly per week, posts "spam" content, such as links to outside articles, and uses a scraped photo from the Internet (Ferrara, 2017).

@TwitchyTeam is considered a "news bot" in this study because it possesses a bot score of 0.70 or higher. The user has a 0.0 percent retweet ratio and mainly posts original tweets with links to articles outside of Twitter. Based on the Twitter account's biography, the user joined Twitter in 2012 and has 71,436 tweets, 237,180 followers and follows 3,692 users. The bot also tweets 400 times per week, on average, according to Hoaxy's Botometer API. Many of these characteristics imply the user is a bot because the user tweets rapidly per week and posts "spam" content (Ferrara, 2017).

@yceek is considered a "news bot" in this study because it possesses a bot score of .70 or higher. The user has a 16 percent retweet ratio and mainly posts original tweets with links to articles outside of Twitter. The user also has a scraped photo from the Internet. Based on the Twitter account's biography, the user joined Twitter in 2016 and has 349,765 tweets, 20,009 followers and follows 16,734 users. The bot also tweets 6,500 times per week, on average, according to Hoaxy's Botometer API. Many of these characteristics imply the user is a bot because the user tweets rapidly per week, regularly posts "spam" content and uses a scraped photo from the Internet (Ferrara, 2017).

@hrkbenowen is considered a "news bot" in this study because it possesses a bot score of 0.70 or higher. The user has a 13 percent retweet ratio and mainly posts original tweets with links to articles outside of Twitter. The user also has a scraped photo from the Internet. Based on the Twitter account's biography, it is a Trump supporter. The user joined Twitter in 2012 and has 213,216 tweets, 67,205 followers and follows 23,993 users. The bot also tweets 1,100 times per week, on average, according to Hoaxy's Botometer API. Many of these characteristics imply the user is a bot because the user tweets rapidly per week, regularly posts "spam" content and uses a scraped photo from the Internet (Ferrara, 2017).

@qkode is considered a "news bot" in this study because it possesses a bot score of .70 or higher. The user has an 11 percent retweet ratio and mainly posts original tweets with links to articles outside of Twitter. The user also has a scraped photo from the Internet. Based on the Twitter account's biography, the user joined Twitter in 2012 and has 364,784 tweets, 21,240 followers and follows 17,865 users. The bot also tweets 7,000 times per week, on average, according to Hoaxy's Botometer API. Many of these characteristics imply the user is a bot because the user tweets rapidly per week, regularly posts "spam" content and uses a scraped photo from the Internet (Ferrara, 2017).

@Alt_Right_ is considered a bot in this study because it possesses a bot score .70 or higher during the study period. This bot was "suspended" from Twitter and further characteristics cannot be retrieved. The bot was most likely removed because it was caught for possessing bot characteristics and posting spam content.

Of the 139 articles (out of 230 top "claim" misinformation articles) the bot users (n = 9) were involved in spreading from January 1, 2017, to December 31, 2017, they all received low engagement. This suggests that Twitter bots aren't always influential; rather they tend to post "spam" content with links to articles from outside sources (See Table 9.)

Table 9: Total amount of articles bots (n = 9) spread and engagement received from the spread of those articles

Bot User Handle	# Misinformation Articles Spread	Total Number of Engagement From All Misinformation Articles Spread
@JoeFreedomLove	61	383 Tweets
@_Breitbot_	9	242 Tweets
@AIIAmericanGirI	24	264 Tweets
@TroyCoby	4	11 Tweets
@TwitchyTeam	8	1008 Tweets
@yceek	8	19 Tweets
@Alt_Right_	3	10 Tweets
@hrkbenowen	16	250 Tweets
@qkoke	6	16 Tweets

Source: Botometer

Discussion

Trends, tone, and engagement of misinformation and fact-checked articles

Previous literature suggests that social media has allowed misinformation, disinformation and fake news to spread more rapidly than ever before (Erlanger, 2017; Vargo, et al., 2016; Waldrop, 2017). The current study found that articles containing misinformation, disinformation, or misleading information spread quickly on Twitter from January 2017 to December 2017, during the first year of Trump's presidency. The majority of misinformation articles that were spread on Twitter received less than 30,000 in engagement, including retweets, mentions, replies, and quotes, while only 10 articles of the total 240 examined received over 30,000 in engagement. This finding suggests that Twitter has created a landscape where misinformation, disinformation, and fake news stories spread between people quickly, but they do not necessarily always receive more engagement than articles posted on Twitter by reliable news sources. For this study, claim "misinformation" articles received much more engagement on Twitter than the fact-checked articles did, which suggests that users and bots were more inclined to post misinformation.

Previous literature suggests that certain types of misinformation may be more consequential than others in terms of the negative impacts on society (Southwell et al., 2015). This study found that different misinformation sources published stories with different tones, which may change how the public perceived a specific person, event, or political view. Interestingly, a majority of the misinformation articles examined from January 2017 to December 2017 had a negative tone. For example, most of the misinformation article topics were related to general politics and Trump. Of those articles published about Trump, most of them were written with a negative tone and published by Politicus USA. On the other hand, the misinformation articles about Trump with a positive tone mostly came from Breitbart and Infowars. This finding suggests that Politicus USA wrote articles in opposition to Trump and published them on social media to manipulate, persuade, or inform the public about negative perspectives about Trump rather than provide objective information. Breitbart and Infowars published mostly positive stories related to Trump, which suggests that they supported his presidency and wrote misinformation articles to persuade the public to support him or oppose those who disagreed with Trump's policies.

Twitter bots spread similar misinformation article topics

Less research has focused on how social bots affect the spread of misinformation articles. Previous research suggests that bots are more likely to speed up the discovery process of content on the Internet and dissemination of particular stories on social media (Manjoo, 2017) and spread information during political campaigns in favor of particular candidates (Conover et al., 2011). In this study, a majority of the users were likely-not-bot users and non-bot users, while a small percentage were bots with a bot score of .70 or higher. The likely-not-bot users in this study had a bot score between .50-.69 and were not considered "bots" in this study because the bot scores did not remain at .70 or higher. In total, nine social bots were involved in the spread of 139 of the top 230 "claim" misinformation articles discovered using the Hoaxy Network API, which suggests that bots are inclined to spread articles containing misinformation, disinformation, or misleading information.

The majority of bots in this study were more likely to spread misinformation on topics from Breitbart, which was the most popular misinformation source, and spread topics related to Trump. Bots will find all the tweets that match the criteria, will perform actions based on the data they find and will be automatically initialized to auto-run on Twitter (Agarwal, 2017). For example, one bot user, *@_breitbot_,* only spread articles from Breitbart, and the most common article topic from January 2017 to December 2017 was about Trump. The most popular bot during the study, *@JoeFreedomLove*, was involved in the spread of the most misinformation articles out of all bot users examined. The majority of articles *@JoeFreedomLove* engaged with on Twitter (through original tweets, retweets, quotes, replies, and mentions) was about general politics from Breitbart.

@AIIAmericanGirI and *@hrkbenowem* are "conservatives" and "Trump supporters," based on their Twitter biographies. The majority of misinformation articles *@AIIAmericanGirI* and *@hrkbenowem* spread on Twitter were from Breitbart about Trump. This finding means the bots were created to automatically publish articles from misinformation sources that wrote "positive" articles about Trump. This finding suggests that these bots were programmed to find articles from specific misinformation sources that matched certain keywords or phrases. From this current study, this finding shows that bot users do affect how misinformation articles are

spread on Twitter, because a majority of them were involved in the spread of similar topics from similar misinformation sources. For example, *@JoeFreedomLove* only received 383 total engagement for all 61 articles the user spread during the study period. This means that bot users are more inclined to spread misinformation articles at a rapid pace, but they are less likely to receive retweets, quotes, replies, and mentions from other users after posting them to Twitter. A majority of the bots examined in this study published stories related to "Trump" and "general politics" from Breitbart, which happened to be the most popular misinformation article topic and source during the study period. Overall, these bots all published mostly political misinformation stories, which suggest they were created to push political propaganda on Twitter. This finding suggests that bots are created to persuade other users on current political topics, campaigns, and candidates by pushing out content quickly. The current study shows that even though bot users do impact the spread of misinformation articles on Twitter, most bots are not very influential.

Limitations and Future Research

A limitation of Botometer is that the collection of bot scores fluctuates and only the most recent 200 tweets and 100 mentions of each user are captured because it tracks current behavioral patterns of users in real time (Mittos, et al., 2018; Shao et al., 2017). Like any technology, Botometer is not flawless, but it has been used successfully in past research to identify likely bots on Twitter (Chavoshi, et al., 2016; Ferrara, et al., 2017; Ferrara 2017; Mittos, et al., 2018; Shao, et al., 2017; Shao et al., 2016; Varol, et al. 2017; Velázquez, et al. 2017). To determine if a user was a bot, manual investigation of each user on Twitter was necessary to match the users with possible bot characteristics. This step ensured that the bot scores were accurate.

A limitation of Hoaxy is that it describes misinformation articles as “claims” because the researchers have stated that they cannot prove that this volume of articles contains inaccurate or unverified information. Rather, Hoaxy relies on outside fact-checking organizations, and it is up to researchers to evaluate the evidence about a claim (Shao et al., 2016). Another limitation of Hoaxy is that not all of the “claim” misinformation articles are connected to a specific “fact-checking” article, according to an interview the researcher of this study conducted with Fil Menczer, a co-creator of Hoaxy and Botometer. The current study relied on Hoaxy as secondary research. Using Hoaxy as the source for “claim” misinformation articles limited the study to the articles collected in Hoaxy’s database. This means that the top 240 “claim” misinformation articles in 2017 were not necessarily the “top” misinformation articles that spread on Twitter that year. This is because Hoaxy may not have included all “claim” misinformation sources in its database. Nonetheless, using just Hoaxy data did reveal trends in misinformation articles and the effect that social bots had on the spread of those articles.

When using the Hoaxy Network API to collect data on how each individual article spread from user to user on Twitter, it was difficult to extract how a tweet spread from one user to another through either a retweet, reply, quote or mention without additional programming. Manual collection was completed with a smaller sample size (of bots and articles) to find general trends in article topics, which is mentioned in the findings and discussion. The manual collection of data ensured greater accuracy but limited the study findings to the “total engagement” for all articles spread. Many of the features associated with the Hoaxy Network API required additional programming to discover the exact article every user was involved in spreading. Hoaxy data only date back to the end of 2016, so article information was not available to collect for comparison from previous years.

Future research could examine how social bot networks affect the spread of misinformation articles or articles containing misinformation from reliable news outlets. Less research has been conducted to examine how social bot networks interact with specific article trends, such as current environmental issues or political scandals. A case study to determine which bots were involved in the spread of top articles containing misinformation around a specific topic may reveal how social bots interact with current misinformation article topics. Less research has focused on what is being done to fight back against misinformation and social bots spreading on social media. Since misinformation and fake news manipulation is a prominent issue in society, a study to examine how social media platforms are countering the spread of inaccurate information and social bots may help determine the underlying problem at large.

Conclusion

Little research has focused on how social bots impact the spread of misinformation articles on Twitter. This study focused on utilizing secondary data from Hoaxy, a platform for tracking how misinformation articles spread on social media, and Botometer, which tracks the likelihood that a Twitter account is a social bot. Using Hoaxy's API features, this study collected top "claim" misinformation articles from Hoaxy's database and the top "most active spreaders" of misinformation articles from January 1, 2017, to December 31, 2017. This research focused on determining how much of an impact social bots had on the spread of misinformation articles during the first year of Trump's presidency compared to non-bot users. Though social bots have an impact on the spread of misinformation articles, this study found that non-bot users were involved in the spread of more misinformation articles because the social bots tend to post "spam" content from outside sources. This finding was important because misinformation, fake news, and disinformation have become more common with the growth of social media and the increase in automated user accounts that spread information rapidly.

Article topics for “misinformation” and “fact-checking” were determined through a quantitative approach, since article topics are constantly changing. The study findings indicate that top misinformation articles in 2017 spread more rapidly on social media than fact-checking articles did. The top misinformation article topics and fact-checking article topics were similar and included information about Trump and politics. The majority of bots in this study were also more likely to spread misinformation on topics from Breitbart, which was the most popular misinformation source, about topics related to Trump and general politics. This finding suggests that bots are more likely to spread articles based on top misinformation trends. Bots that are correlated with a specific political party or worldview are also more likely to spread misinformation articles based on topics of interest.

Most important, this study suggests that with the ever-growing social media universe, bots are affecting the way that news is spread, understood, and contextualized in society. A majority of misinformation article topics in 2017 were negative and related to politics, which suggests that misinformation articles were created to manipulate public opinion and spread subjective information. With the increase in technology and capacity for misinformation, disinformation and fake news to spread quicker than ever before, bots are programmed to rapidly publish misleading articles based on current “news” and “misinformation” trends.

Appendix A: Codebook protocol for misinformation and fact-checking articles

CONTENT VARIABLES:

VAR 01: "Claim" Misinformation Article Title
This refers to the specific article title (full name of the misinformation or fact-checking article).
Example: CBS Confirms Trump Brought His Own Studio Audience To Clap For Him At CIA Speech

VAR 02: "Claim" Misinformation Source
This refers to the misinformation source the article was published by.

1. Breitbrat
2. Infowars
3. PoliticusUSA
4. Your News Wire
5. Before It's News
6. Big Partisan Report
7. Twitchy
8. The Blaze
9. The Onion
10. US Uncut
11. Occupy Democrats
12. WND
13. ANO News

VAR 03: Claim "Misinformation" or "Fact-Checking" Article
This refers to if the article was either from a misinformation (claim) source or fact-checking source. *Example: PoliticusUSA (0) or Snopes (1).*

0. Claim
1. Misinformation

VAR 04: Number of Tweets
This refers to the amount of engagement the misinformation article received each month
0. 0 – 30,000
1. 1 – 30,000 to highest

VAR 05: Article Topic

The article topics were determined quantitatively. Since this study monitored misinformation article topics and fact-checking article trends from January 1, 2017 to December 31, 2017, the article topics were created **while** examining the headlines, keywords and content within each article. Article topics about Obama, Russia, immigration, the military and other were later placed in the "general politics" category.

1. Crime
2. Entertainment
3. Immigration
4. Military
5. Obama
6. Trump
7. General Politics
8. Satire
9. Race
10. Russia
11. Women's Rights/Feminism
12. Conspiracy Theory
13. Terrorism
14. The Media
15. Other

VAR 06: Tone

This refers to the tone of the misinformation article, which includes whether the headline and content of the article were in favor of, against or neutral about a topic or individual.
Positive tone – contains favorable language, such as strong support and excitement about a political candidate, event or individual.
Negative tone – contains a bias headline, include information in support of one side of an issue or include stereotypical or exaggerated information.
Neutral tone – selected if the article wasn't clearly positive or negative.

1. Positive
2. Negative
3. Neutral

VAR 07: Month posted on Twitter
This refers to the month the article was first published on Twitter.

1. January
2. February
3. March
4. April
5. May
6. June
7. July
8. August
9. September
10. October
11. November
12. December

VAR 08: Month Monitored
This refers to the month the article was monitored on Twitter during the study period.

1. January
2. February
3. March
4. April
5. May
6. June
7. July
8. August
9. September
10. October
11. November
12. December

VAR 09: URL
The original URL of the articles (misinformation or fact-check)
Example: http://www.politicususa.com/2017/01/23/cbs-confirms-reports-donald-trump-brought-cheering-props-cia-visit.html

Appendix B: Codebook protocol for "top users"

VAR 01: User Handle	This refers to the name of the Twitter user Example: @TheRealDonaldTrump
VAR 02: Number of Tweets	This refers to the amount of engagement the user received that month.
VAR 03: Bot Score	This section refers to whether the user was a bot or not. Bot scores were determined using Hoaxy's Botometer API. If the bot score of a user was .70 or higher, it was considered a bot for this study. 1 (0-.49) 2 (.50-.59) 3 (.70-.99)
VAR 04: Date Monitored	This section refers to when the Twitter user was monitored on Twitter during the study period. 1 (January – February 2017) 2 (February – March 2017) 3 (March – April 2017) 4 (April – May 2017) 5 (May – June 2017) 6 (June – July 2017) 7 (July – August 2017) 8 (August – September 2017) 9 (September – October 2017) 10 (October – November 2017) 11 (November – December 2017) 12 (December 2017)
VAR 06: Spread Type	This section refers to if the user was involved in spreading claim or fact-checking articles. 0 (claim) 1 (fact-check)

References

Abokhodair, N., Yoo, D., & McDonald, D. W. (2015, February). Dissecting a social botnet: Growth, content and influence in Twitter. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 839-851). ACM.

Agarwal, A. (2017, July 21). How to write a Twitter bot in 5 minutes. *Digital Inspiration.* Retrieved from https://www.labnol.org/internet/write-twitter-bot/27902/

Aiello, L. M., Deplano, M., Schifanella, R., & Ruffo, G. (2012). People are strange when you're a stranger: Impact and influence of bots on social networks. *arXiv.* Retrieved from https://arxiv.org/pdf/1407.8134.pdf

Albright, J. (2017). Welcome to the era of fake news. *Media and Communication, 5*(2), 87. doi:10.17645/mac.v5i2.977

Altheide, D. & Schneider, C. (2013). *Qualitative media analysis* 55 City Road, London: SAGE Publications Ltd doi: 10.4135/9781452270043

Balmas, M. (2012). When fake news becomes real. *Communication Research, 41*(3), 430-454. doi:10.1177/0093650212453600

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. presidential election online discussion. *First Monday, 21*(11). doi:10.5210/fm.v21i11.7090

Bilton, R. (2017, June 22). The latest class of Knight Prototype projects will tackle trust, media literacy, and yes, fake news. *Nieman Lab*. Retrieved from http://www.niemanlab.org/2017/06/the-latest-class-of-knight-prototype-projects-will-tackle-trust-media-literacy-and-yes-fake-news/

Bode, L., & Vraga, E. K. (2015). In related news, that was wrong: The correction of misinformation through related stories functionality in social media. *Journal of Communication*, *65*(4), 619-638.

Borchers, C. (2017, October 26). Trump falsely claims (again) that he coined the term 'fake news'. *The Washington Post.* Retrieved from https://www.washingtonpost.com/news/the-fix/wp/2017/10/26/trump-falsely-claims-again-that-he-coined-the-term-fake-news/?utm_term=.8b321463011a

Chavoshi, N., Hamooni, H., & Mueen, A. (2016, November). Identifying correlated bots in Twitter. *International Conference on Social Informatics*, 14-21. doi:10.1007/978-3-319-47874-6_2

Chu, Z., Gianvecchio, S., Wang, H., & Jajodia, S. (2012). Detecting automation of Twitter accounts: Are you a human, bot, or cyborg? *IEEE Transactions on Dependable and Secure Computing, 9*(6), 811-824. doi:10.1109/tdsc.2012.75

Ciampaglia, G. L. (2017). Fighting fake news: a role for computational social science in the fight against digital misinformation. *Journal of Computational Social Science, 1*(1), 147-153. doi:10.1007/s42001-017-0005-6

Conover, M. D., Gonçalves, B., Ratkiewicz, J., Flammini, A., & Menczer, F. (2011, October). Predicting the political alignment of twitter users. *Privacy, Security, Risk and Trust (PASSAT),* 192-199.

Davis, C. A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. (2016, April). Botornot: A system to evaluate social bots. *Proceedings of the 25th International Conference Companion on World Wide Web*, 273-274.

Desigaud, C., Howard, P., Bradshaw, S., Kollanyi, B., Bolsover, G. (2017). Junk news and bots during the French presidential election: What are French voters sharing over Twitter? *Comprop Data Memo.* Retrieved from http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/04/What-Are-French-Voters-Sharing-Over-Twitter-v9.pdf

Dewey, C. (2016, October 19). One in four debate tweets comes from a bot. Here's how to spot them. *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/the-intersect/wp/2016/10/19/one-in-four-debate-tweets-comes-from-a-bot-heres-how-to-spot-them/?utm_term=.c9950bd061e6

Diakopoulos, N., & Koliska, M. (2016). Algorithmic transparency in the news media. *Digital Journalism, 5*(7), 809-828. doi:10.1080/21670811.2016.1208053

Dörr, K. N., & Hollnbuchner, K. (2016). Ethical challenges of algorithmic journalism. *Digital Journalism, 5*(4), 404-419. doi:10.1080/21670811.2016.1167612

Ecker, U., Lewandowsky, S., Chang, P., & Pillai, R. (2014). Supplemental material for the effects of subtle misinformation in news headlines. *Journal of Experimental Psychology:* Applied *20*(4), 323-335. doi:10.1037/xap0000028

Erlanger, S. (2017, December 12). 'Fake news,' Trump's obsession, is now a cudgel for strongmen*.* *The New York Times*. Retrieved from https://www.nytimes.com/2017/12/12/world/europe/trump-fake-news-dictators.html?smprod=nytcore-ipad&smid=nytcore-ipad-share&_r=0

Faris, R., Roberts, H., Etling, B., Bourassa, N., Zuckerman, E., Benkler, Y. (2017). Partisanship, propaganda, and disinformation: Online media and the 2016 U.S. presidential election. *Berkman Klein Center Research Publication.*

Ferrara, E. (2017). Disinformation and social bot operations in the run up to the 2017 French presidential election. *First Monday, 22*(8), 1-33. doi:10.5210/fm.v22i8.8005

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), 96-104. doi:10.1145/2818717

Geiger, R. S. (2014). Bots, bespoke, code and the materiality of software platforms. *Information, Communication & Society, 17*(3), 342-356. doi:10.1080/1369118x.2013.873069

Geiger, R. S. (2016). Bot-based collective blocklists in Twitter: the counterpublic moderation of harassment in a networked public space. *Information, Communication & Society, 19*(6), 787-803. doi:10.1080/1369118x.2016.1153700

Gottfried, J., & Shearer, E. (2016, May 26). News use across social media platforms 2016. *Pew Research Center*. Retrieved from http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/

Griffin, A. (2014, August 18). Facebook to tag satirical articles to stop users falling for the Onion's jokes. *The Independent*. Retrieved from https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-to-flag-satirical-articles-to-stop-users-falling-for-the-onions-jokes-9674483.html

Hopmann, D. N., Vliegenthart, R., Vreese, C. D., & Albæk, E. (2010). Effects of election news coverage: How visibility and tone influence party choice. *Political Communication, 27*(4), 389-405. doi:10.1080/10584609.2010.516798

Karlova, N. A., & Fisher, K. E. (2013). A social diffusion model of misinformation and disinformation for understanding human information behaviour. *Information Research*, *18*(1), 1-17.

Kew, B. (2017, March 10). Major Impact: president Donald Trump's first 50 days in office. *Breitbart*. Retrieved from http://www.breitbart.com/big-government/2017/03/10/major-impact-president-donald-trumps-first-50-days-in-office/

Kollanyi, B. (2016). Where do bots come from? An analysis of bot codes shared on GitHub. *International Journal of Communication*, 10, 4932-4951.

Lokot, T., & Diakopoulos, N. (2015). News bots. *Digital Journalism, 4*(6), 682-699. doi:10.1080/21670811.2015.1081822

Lu, J. (2014, August 28). Is anyone actually reading your tweets? *The Washington Post.* Retrieved from https://www.washingtonpost.com/news/technology/wp/2014/08/28/is-anyone-actually-reading-your-tweets/?utm_term=.3eb698a88dae

Manjoo, F. (2017, May 31). How Twitter is being gamed to feed misinformation. *The New York Times*. Retrieved from https://www.nytimes.com/2017/05/31/technology/how-twitter-is-being-gamed-to-feed-misinformation.html?_r=0

Marchi, R. (2012). With Facebook, blogs, and fake news, teens reject journalistic "objectivity". *Journal of Communication Inquiry, 36*(3), 246-262. doi:10.1177/0196859912458700

Mittos, A., Blackburn, J., & De Cristofaro, E. (2018). 23andMe confirms: I'm super white. analyzing Twitter discourse on genetic testing. *arXiv*. Retrieved from https://arxiv.org/pdf/1801.09946.pdf

Montal, T., & Reich, Z. (2016). I, robot. You, journalist. Who is the author? *Digital Journalism, 5*(7), 829-849. doi:10.1080/21670811.2016.1209083

Mullin, B. (2017, June 22). Knight Foundation awards $1 million to projects aimed at fighting misinformation. *Poynter.* Retrieved from https://www.poynter.org/news/knight-foundation-awards-1-million-projects-aimed-fighting-misinformation

Napoli, P. M. (2014). Automated Media: An Institutional Theory Perspective on Algorithmic Media Production and Consumption. *Communication Theory, 24*(3), 340-360. doi:10.1111/comt.12039

Oentaryo, R. J., Murdopo, A., Prasetyo, P. K., & Lim, E. (2016). On profiling bots in social media. *Computer Science Social Informatics,* 92-109. doi:10.1007/978-3-319-47880-7_6

Rannard, G. (2017, December 31). How fake news plagued 2017. *BBC News.* Retrieved from http://www.bbc.com/news/world-42487425

Reaney, P. (2016, December 21). U.S. university launches tool to show how fake news spreads. *Reuters.* Retrieved from https://www.reuters.com/article/us-usa-tech-fake-idUSKBN14A2CO

Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & Menczer, F. (2017). The spread of fake news by social bots. *ArXiv.* Retrieved from https://arxiv.org/pdf/1707.07592.pdf

Shao, C., Ciampaglia, G. L., Flammini, A., & Menczer, F. (2016, April). Hoaxy: A platform for tracking online misinformation. *Proceedings of the 25th international conference companion on world wide web*, 745-750.

Shao, C., Hui, P. M., Wang, L., Jiang, X., Flammini, A., Menczer, F., & Ciampaglia, G. L. (2018). Anatomy of an online misinformation network. *arXiv.* Retrieved from https://arxiv.org/pdf/1801.06122.pdf

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media. *ACM SIGKDD Explorations Newsletter,19*(1), 22-36. doi:10.1145/3137597. 3137600

Silverman, C. (2016). This analysis shows how viral fake election news stories outperformed real news on Facebook. Retrieved from https://www.buzzfeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook

Solon, O. (2017, December 18). How Syria's White Helmets became victims of an online propaganda machine. Retrieved from https://www.theguardian.com/world/2017 /dec/18/syria-white-helmets-conspiracy-theories

Southwell, B. G., & Thorson, E. A. (2015). The prevalence, consequence, and remedy of misinformation in mass media systems. *Journal of Communication*. 65(4). 589-595.

Stieglitz, S., Brachten, F., Ross, B., & Jung, A. K. (2017). Do social bots dream of electric sheep? A categorisation of social media bot Accounts. *arXiv.* Retrieved from https://arxiv.org/pdf/1710.04044.pdf

Tam, P. (2016, December 27). Daily report: inside Snopes, a fighter of fake news. *The New York Times*. Retrieved from https://www.nytimes.com/2016/12/27/technology/daily-report-inside-snopes-a-fighter-of-fake-news.html?mtrref=undefined&gwh=2B97 E0658E535827B87AE359CC078D0B&gwt=pay

Tan, A. S., Lee, C. J., & Chae, J. (2015). Exposure to health (mis) information: Lagged effects on young adults' health behaviors and potential pathways. *Journal of Communication,* 65(4), 674-698.

Tudjman, M., & Mikelic, N. (2003). Information science: Science about information, misinformation and disinformation. *Proceedings of Informing Science + Information Technology Education*, 1513-1527.

Van Dalen, A. (2012). The algorithms behind the headlines: How machine-written news redefines the core skills of human journalists. *Journalism Practice*, *6*(5-6), 648-658.

Vargo, C. J., Guo, L., & Amazeen, M. A. (2017). The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. *New Media & Society, 20*(5), 2028-2049. doi:10.1177/1461444817712086

Velázquez, E., Yazdani, M., & Suárez-Serrato, P. (2017). Socialbots supporting human rights. *arXiv*. Retrieved from https://arxiv.org/pdf/1710.11346.pdf

Waldrop, M. (2017). News Feature: The genuine problem of fake news. *Proceedings of the National Academy of Sciences, 114*(48),12631–12634. doi:10.1073/pnas.1721073115

Weeks, B. E. (2015). Emotions, partisanship, and misperceptions: How anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. Journal of Communication, 65(4), 699-719.

Woolley, S. C. (2016). Automating power: Social bot interference in global politics. *First Monday, 21*(4). doi:10.5210/fm.v21i4.6161

Woolley, S. C., & Howard, P. N. (2016). Political communication, computational propaganda, and autonomous Agents. *International Journal of Communication*, *10*(9), 4882–4890.

Yin, R. K. (2017). *Case study research and applications: Design and methods*. Thousand Oaks, CA: SAGE Publications.

Zhao, S. (2006). Humanoid social robots as a medium of communication. *New Media & Society,* 8(3), 401-419. doi:10.1177/1461444806061951