Over Half of Tweets about E-Cigs Might Come from Bots

Executive Summary

Among youth, the most popular method of nicotine consumption is currently through the use of electronic cigarettes, or e-cigarettes. Rates of e-cigarette use have climbed to staggering levels over the past decade; nationally, e-cigarette use has risen sharply among high school students, from 1.5% in 2011 to 20.8% in 2018. Use among middle school students has also increased, from 0.6% in 2011 to 4.9% in 2018.[1,2] The e-cigarette market continues to innovate and expand. Online sales make up a substantial portion of the overall market for both adults and adolescents, with estimates that over one-third of teens who use e-cigarettes purchase them either in a store or online.[3] Despite the fact that each state has its own laws restricting sales of e-cigarettes to minors, two recent studies have shown that between 60% and 98% of e-cigarette sales transactions do not employ reliable age verification techniques.[4,5] Additionally, there are many aspects of e-cigarettes that are not widely understood by the general public, including whether e-cigarettes are safe, if there is evidence of e-cigarette use as an effective strategy for helping people quit smoking, and the likelihood of e-cigarette users progressing to cigarette and other drug use. Gaps in knowledge, as well as the emerging news about the negative health impacts of e-cigarette use, make the topic one that is particularly susceptible to misinformation.

---

[1] Gentzke, A. S., Creamer, M., Cullen, K. A., Ambrose, B. K., Willis, G., Jamal, A., & King, B. A. (2019). Vital signs: Tobacco product use among Middle and High School Students—United States, 2011–2018. *Morbidity and Mortality Weekly Report, 68*(6), 157.

[2] Cullen, K. A., Ambrose, B. K., Gentzke, A. S., Apelberg, B. J., Jamal, A., & King, B. A. (2018). Notes from the field: Use of electronic cigarettes and any tobacco product among middle and high school students—United States, 2011–2018. *Morbidity and Mortality Weekly Report, 67*(45), 1276.

[3] Pepper, J. K., Coats, E. M., Nonnemaker, J. M., & Loomis, B. R. (2019). How do adolescents get their e-cigarettes and other electronic vaping devices? *American Journal of Health Promotion, 33*(3), 420–429. [0890117118790366].

[4] Williams, R. S., Derrick, J., Liebman, A. K., LaFleur, K., & Ribisl, K. M. (2018). Content analysis of age verification, purchase and delivery methods of internet e-cigarette vendors, 2013 and 2014. *Tobacco Control, 27*(3), 287–293.

[5] Williams, R. S., Derrick, J., & Ribisl, K. M. (2015). Electronic cigarette sales to minors via the internet. *JAMA Pediatrics, 169*(3), e156.

A study conducted by the public health nonprofit The Public Good Projects (PGP) with funding from The Nicholson Foundation discovered that over half of all messages transmitted through public media sources in the United States regarding e-cigarettes and tobacco products may be posted by automated accounts, or bots. Between February 1, 2019, and June 1, 2019, PGP collected publicly available media data across social media, websites, online videos, print media, and broadcast television. Out of a total national sample of 1,288,378 Twitter messages related to e-cigarettes or tobacco, 22.6% of messages came from likely humans, and 77.4% of messages came from accounts suspected or highly likely to be bots, many of which are dedicated to promoting misinformation about nicotine and e-cigarettes. PGP conducted a large part of the analysis on Twitter, as bots are heavily active on Twitter's platform. Even though everyone may not be on Twitter, marketers and social network researchers widely recognize that Twitter discussions often bleed into other public forums, both offline and online. While there are nuance and complexity around what constitutes a bot, these numbers provide an estimation of who is generating these discussions and build upon previous studies showing similar results. Earlier research has suggested that as many as 70% of tweets about e-cigarettes come from bots posing as real people and that commercially motivated bots may be perpetuating misinformation about the efficacy of e-cigarettes in smoking cessation.[6] However, this is the first study to show the enormous potential scope of the national bot discussion regarding e-cigarettes and tobacco. Approximately 80% of all messages during the study period came from suspected bots on Twitter. Of the top suspected pro-vaping bot accounts active during this period, most appear to be adopting the following strategies: promoting e-cigarette sales, posing as passionate pro-vaping individuals who encourage e-cigarette use, and creating groundswell movements based on misinformation, or a combination of these. While some suspected bots do promote anti-vaping messaging, their number pales in comparison to the number of those promoting pro-vaping messaging. Bots can function individually or in conjunction with one another through "bot networks," or "botnets," which may consist of dozens or hundreds of individual bots working in concert to

---

[6] Allem, J.-P., Ferrara, E., Uppu, S. P., Cruz, T. B., & Unger, J. B. (2017). E-cigarette surveillance with social media data: Social bots, emerging topics, and trends. *JMIR Public Health and Surveillance, 3*(4), e98.

promote the same message. The graphic below shows pro-vaping discussion and the ways that bots and humans interact with and share information about e-cigarettes. Each circle represents a Twitter account. The size of each circle reflects the size of that account's following. The lines between circles represent the sharing of information. The color of a circle indicates the likelihood of that account being a human (white) or bot (light pink to dark pink, with dark pink/ purple meaning it is highly likely to be a bot).
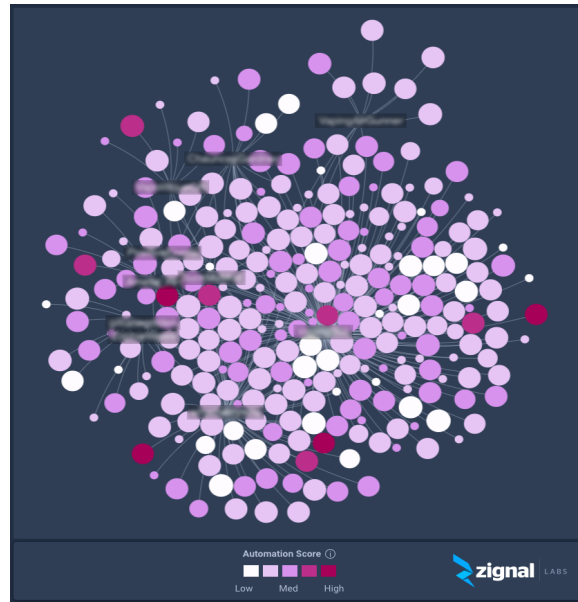
Results suggest that bots are actively promoting vaping as a viable strategy for quitting smoking, and suspected bots commonly referenced the importance of quitting smoking. While this may please public health authorities, a deeper examination shows that the majority of these messages encourage transitioning from cigarettes to e-cigarettes or vaping—even though studies have not confirmed that e-cigarettes are effective as a cessation aid. Suspected bots commonly embed anti-tobacco sentiment within pro-vaping messaging. At face value, these suspected bot messages could appear to be anti–tobacco industry, anti-smoking, and pro-health. However, what appears to be messaging that is adopting anti-tobacco stances, for example highlighting the transgressions of the tobacco industry, could be pro-vaping content generated by bots.

PGP examined conversations from three sources: highly likely bots (accounts that display activity that is highly characteristic of a bot), suspected bots (accounts that display activity that could indicate a bot account), and humans (accounts that have a very high likelihood of being humans). PGP then looked at the proportion of discussion about specific e-cigarette and tobacco products, displayed in the "Conversation by Product" chart in the report below. Of the product groups, e-cigarettes were more frequently mentioned than other tobacco products. This was followed by messages about cigarettes, with other products (such as cigars, hookah, cigarillos and smokeless) showing variation across the three groups.

Themes, or how messages were framed, showed similarities between humans and suspected bots. Conversation in these two groups tended to focus on quitting smoking. Accounts that had the highest likelihood of being a bot focused primarily on promoting sales. Conversations about the health effects of e-cigarettes and tobacco were nearly the same across groups, with accounts highly likely to be bots showing a slightly higher proportion of messages mentioning youth.

Results from this study confirm not only the ease of finding e-cigarette advertising but also the high volume of this advertising relative to all other messages regarding e-cigarettes on one of the world's largest and most influential social media platforms. However, questions remain regarding who is behind these suspected bots. Are companies themselves funding botnets, or are the online marketing activities promoting e-cigarettes generated by individuals who run online stores, advertisers, individual e-cigarette advocates, or a combination of each of these? And whoever they are, where are they physically based? PGP plans on pursuing these questions in its next round of research. In the meantime, the nonprofit organization does not recommend that regulators or others attempt to shut down suspected bot accounts. Rather, public health officials should focus on monitoring tobacco product advertising and messaging and using this information to create more effective communication strategies. Bots themselves are not good or bad, but understanding their messaging and role in a dynamic and fast-changing media landscape is critical for public health.

Table of Contents

Introduction

The Public Good Projects (PGP) is collaborating with The Nicholson Foundation to examine publicly available media data around e-cigarettes and tobacco products to examine the prevalence and influence of online automated, or bot, networks. To understand the types of conversations that are occurring online around e-cigarettes and other tobacco products, PGP conducted a systematic analysis of media data to determine the following:

- the amount of discourse around e-cigarettes and other tobacco products;
- the role of automated accounts, or bots, in conversation; and
- common conversation themes that bots are promoting.

Electronic Cigarettes

Among youth, the most popular method of nicotine consumption is currently through the use of electronic cigarettes, or e-cigarettes. E-cigarettes are known by various names, including electronic nicotine delivery systems, e-cigs, vapes, vape pens, e-cigars, hookah pens, tanks, and mods. Regardless of the name, e-cigarettes are battery-operated devices used to inhale an aerosol that typically contains flavorings, nicotine, and other chemicals, such as propylene glycol or glycerin.

Rates of e-cigarette use have experienced a staggering climb over the past decade. Among adults, the national prevalence of current e-cigarette use was around 3% in 2017, compared to 4.4% in New Jersey.[7,8] Nationally, current e-cigarette use has seen a particularly sharp rise among high schoolers, from 1.5% in 2011 to 20.8% in 2018.[9,10] Correspondingly, use among

---

[7] Wang, T. W., Asman, K., Gentzke, A. S., Cullen, K. A., Holder-Hayes, E., Reyes-Guzman, C., Jamal, A., Neff, L., & King, B. A. (2018). Tobacco product use among adults—United States, 2017. *Morbidity and Mortality Weekly Report, 67*(44), 1225.

[8] Centers for Disease Control and Prevention (CDC). (2017). BRFSS Prevalence & Trends Data. Retrieved June 23, 2019, from https://nccd.cdc.gov/BRFSSPrevalence

[9] Gentzke, A. S., Creamer, M., Cullen, K. A., Ambrose, B. K., Willis, G., Jamal, A., & King, B. A. (2019). Vital signs: Tobacco product use among Middle and High School Students—United States, 2011–2018. *Morbidity and Mortality Weekly Report, 68*(6), 157.

middle schoolers has also increased, from 0.6% in 2011 to 4.9% in 2018. In New Jersey, use among high schoolers has increased from 6.1% in 2012 to 9.6% in 2016.[11,12] However, given the recent rise in e-cigarette use across the country (and the fact that JUUL joined the market in 2017), it is likely that the current (2019) rates of e-cigarette use are substantially higher.

Although they have been widely used for over a decade, there are many aspects of e-cigarettes that are still relatively misunderstood, including safety, efficacy as a cessation aid, and likelihood of progression from e-cigarette to cigarette use (particularly among youth). Evidence suggests that e-cigarettes may be relatively less harmful than regular cigarettes because some appear to contain fewer toxic chemicals and do not burn tobacco leaves.[13] Some researchers have claimed that e-cigarettes could potentially be beneficial for adult smokers who are trying to quit smoking, but further study is necessary to confirm whether it is effective as a cessation aid.[14] As e-cigarettes can potentially be both harmful and beneficial and the scientific evidence base is still growing, there is a general lack of awareness and information about e-cigarettes.[15,16] For example, a recent study reported that the majority of e-cigarette-using teens were unaware of whether their e-cigarette contains nicotine.[17] In addition to the lack of information among the general public, there are clear knowledge gaps

---

[10] Cullen, K. A., Ambrose, B. K., Gentzke, A. S., Apelberg, B. J., Jamal, A., & King, B. A. (2018). Notes from the field: Use of electronic cigarettes and any tobacco product among middle and high school students—United States, 2011–2018. *Morbidity and Mortality Weekly Report*, 67(45), 1276.

[11] New Jersey Department of Health. (2016). 2016 NJ Youth Tobacco Survey. Retrieved March 27, 2019, from https://www.nj.gov/health/fhs/tobacco/documents/NJYTS%202016%20State%20Report_Final.pdf

[12] New Jersey Department of Health. (2015). 2014 NJ Youth Tobacco Survey. Retrieved March 27, 2019 from https://nj.gov/health/fhs/tobacco/documents/nj_yts_report_final_120916.pdf

[13] Centers for Disease Control and Prevention. (2018). About Electronic Cigarettes (E-Cigarettes). Retrieved February 6, 2019, from https://www.cdc.gov/tobacco/basic_information/e-cigarettes/about-e-cigarettes.html

[14] Ibid.

[15] Coleman, B. N., Johnson, S. E., Tessman, G. K., Tworek, C., Alexander, J., Dickinson, D. M., Rath, J., Green, K. M. (2016). "It's not smoke. It's not tar. It's not 4000 chemicals. Case closed": Exploring attitudes, beliefs, and perceived social norms of e-cigarette use among adult users. *Drug and Alcohol Dependence, 159*, 80–85. https://doi.org/https://doi.org/10.1016/j.drugalcdep.2015.11.028

[16] Webb Hooper, M., & Kolar, S. K. (2017). Racial/ethnic differences in electronic cigarette knowledge, social norms, and risk perceptions among current and former smokers. *Addictive Behaviors, 67*, 86–91. https://doi.org/https://doi.org/10.1016/j.addbeh.2016.12.013

[17] Alexander, J. P., Williams, P., & Lee, Y. O. (2018). Youth who use e-cigarettes regularly: A qualitative study of behavior, attitudes, and familial norms. *Preventive Medicine Reports, 13*, 93–97. https://doi.org/10.1016/j.pmedr.2018.11.011

among those who work in health fields. Healthcare providers have reported having insufficient guidelines for advising their patients about e-cigarettes.[18] These gaps in knowledge make e-cigarettes a topic that is particularly susceptible to misinformation. As more adults and teens continue to use e-cigarettes, it is critical to understand the types of information—and misinformation—they are exposed to and which types of accounts are circulating messages.

The Role of Bots

Researchers have only begun to examine the role of bots in online conversation around health topics. Bots make up a significant portion of internet conversation, with estimates that bots drive over half of all internet traffic.[19] There are a variety of types of bots that are active online, including bots that perform useful functions. These include bots that provide health tips or alert the public to health threats and fraud detection bots that alert customers to potentially fraudulent purchases on a credit card. However, bots may also seek to promote a nefarious agenda, spreading information that is misleading and disrupts public health. Bots can function individually or in conjunction with one another through "bot networks," or "botnets," which may consist of dozens or hundreds of individual bots working in a concerted manner to promote the same message.

As researchers learn more about the ways in which bots can meddle with the information consumed online, it is becoming clear that health topics are a prime target—particularly topics that are hotly contested or lack a strong evidence base, such as e-cigarettes. Although the evidence base for the role of bots related to e-cigarettes is sparse, some preliminary research has suggested that as many as 70% of tweets about e-cigarettes come from bots posing as real people and that commercially motivated bots may be perpetuating misinformation about

---

[18] Kanchustambham, V., Saladi, S., Rodrigues, J., Fernandes, H., Patolia, S., & Santosh, S. (2017). The knowledge, concerns and healthcare practices among physicians regarding electronic cigarettes. *Journal of Community Hospital Internal Medicine Perspectives, 7*(3), 144–150. https://doi.org/10.1080/20009666.2017.1343076

[19] Zignal Labs, Inc. (2018). Infographic: The Social Media Bot Problem. Retrieved June 23, 2019, from https://zignallabs.com/the-social-media-bot-problem/

the efficacy of e-cigarettes in smoking cessation.[20] As the e-cigarette market continues its explosive growth among youth and adults, it is critical to understand the frequency and content of messages promoted by bots. PGP uses its bot detection technologies to understand the ways that bot accounts are driving conversation, pushing specific agendas, and potentially distorting public dialogue concerning key health issues. With this knowledge, researchers can alert the greater public health community to the most common tactics, conversation themes, and types of misinformation most frequently perpetuated and circulated by bots. Armed with this information, public health and healthcare professionals can then respond.

Methods

As a first step in examining bot-driven conversation, PGP monitors and analyzes media data for references to e-cigarette and tobacco use. Data consist of all publicly available data transmitted across multiple media sources, including the following:

- social and digital media, such as Twitter, Reddit, YouTube, Vimeo, online forums, online Q&A websites, review sites;
- online news sites, and blogs;
- print media, such as trade journals, magazines, and newspapers; and
- broadcast television, and many more.

PGP collects all publicly available media data references to e-cigarettes and tobacco using a complex keyword query constructed using Boolean search methodologies (using "AND" and "EXCLUDE" terms). PGP selects words for the keyword query on the basis of peer-reviewed literature and PGP's experience with content analyses. This process produces a vast database of all publicly available conversations related to the topic of interest. In addition to analyzing the amount of conversation around specific e-cigarette and tobacco products, PGP's software

---

[20] Allem, J.-P., Ferrara, E., Uppu, S. P., Cruz, T. B., & Unger, J. B. (2017). E-cigarette surveillance with social media data: Social bots, emerging topics, and trends. *JMIR Public Health and Surveillance, 3*(4), e98.

allows for analysis of posts by content theme. Content themes are synonymous with "message frames," or the ways in which the public conceptualizes a topic. Through previous e-cigarette research that PGP conducted in New Jersey with the New Jersey Prevention Network, PGP has identified several content themes within conversation around e-cigarettes:

Theme	Description
Addiction	References any addiction to tobacco, e-cigarettes, or nicotine.
Health Effects	References long-term negative health effects of tobacco, e-cigarettes, or nicotine, such as lung cancer, secondhand smoke, diabetes, or chronic diseases; short- term negative health effects, such as bad breath or yellowed fingernails; or general risks related to use of products. With reference to conversation around e-cigarettes, the negative health effects of tobacco are commonly referenced as a reason for using e-cigarettes.
Policy	References policies related to e-cigarettes or tobacco or regulations and taxes. Conversation within this theme commonly references changes to the minimum age to purchase e-cigarettes or tobacco or policies around regulation of production and distribution of e-cigarettes.
Sales	References online sales, advertisements, or promotions of e-cigarettes or tobacco or articles about sales of e-cigarettes or tobacco products.
Quitting	References tobacco and nicotine cessation; often references quitting cigarettes through the use of e-cigarettes.
Research	References research related to e-cigarettes or tobacco.
Youth	References use of e-cigarettes or tobacco by youth or education of youth on issues related to e-cigarette or tobacco use (often including educational anti-tobacco campaigns).

PGP is able to identify which posts have a high likelihood of originating from bots and which have a high likelihood of originating from humans. This analysis occurs on Twitter, given that

bots are heavily active on Twitter's platform.[21] PGP researchers examine multiple account characteristics to determine the likelihood of a post being from a bot, including (but not limited to) the frequency and timing of posts, the number of posts, the number of followers, and engagement with other accounts. Bots, particularly those created with malicious intent, are incredibly nuanced and are often designed to appear exactly like a human. Therefore, a simple examination of basic account and profile characteristics is insufficient to gauge the likelihood of automation. PGP examined conversation from three types of sources: likely bots (accounts that have a very high likelihood of being bots); suspected bots (accounts that display activity that could indicate a bot account but cannot be confirmed with certainty); and humans (accounts that have a very high likelihood of being humans). By separating conversation into these groups, PGP can conduct an examination of data that can provide a more holistic or in-depth understanding of discourse.

For the purposes of this report, PGP analyzed bot conversation at the national and New Jersey state levels. PGP included national-level bot activity as many accounts reaching New Jersey residents either are not based in New Jersey or are based in an indeterminate location due to users employing location-shielding apps or merely turning off geolocation. Some have suggested that this may be particularly common among bot accounts, as they have greater incentive to hide their locations.

The data presented below will allow researchers to gather insight from conversation around e-cigarettes as well as other tobacco products (cigarettes, cigars, cigarillos, little cigars, smokeless tobacco, and hookah). Conversation around e-cigarettes often references tobacco, for example in discussions about using e-cigarettes as a method for quitting cigarettes. It is critical to understand the conversation that is occurring related to tobacco because it informs the conversation happening around e-cigarettes. Additionally, with many of the large e-cigarette companies owned by big tobacco, researchers must keep an eye on the entire

---

[21] Spence, S. (2017). How to build your own Twitter bot in less than 30 minutes. Retrieved June 25, 2019, from https://venturebeat.com/2017/02/02/how-to-build-your-own-twitter-bot-in-less-than-30-minutes/

spectrum of tobacco products to gain context on overall developments. With this in mind, the results below place a particular focus on e-cigarettes and highlight results from other tobacco products if they are notable.

Results

Between February 1, 2019, and June 1, 2019, the overall search query resulted in a total sample of 1,288,378 Twitter messages at the national level and 14,143 within New Jersey. These numbers are inclusive of all conversation around tobacco products and e-cigarettes. Nationally, analysis showed that just over 20% of all conversation during the data collection period can be attributed to humans, with approximately 20% attributed to suspected bots and nearly 60% originating from likely bots.

Conversation within New Jersey showed different patterns, with humans appearing to play less of a role in conversation compared to that seen nationally. Just over 10% of overall conversation in New Jersey originates from humans, a further 50% comes from suspected bots, and around 40% originates from likely bots. The tables below present more information on the amount of conversation originating from humans, suspected bots, and likely bots, by geographic location.

United States	Humans	Suspected Bots	Likely Bots
Percentage of Total Conversation	22.6%	20.8%	56.6%

New Jersey	Humans	Suspected Bots	Likely Bots
Percentage of Total Conversation	10.6%	49.7%	39.7%

Analysis by Product

PGP analysts examined the proportion of conversation about specific e-cigarette and tobacco products. Across both geographic areas and all three groups (humans, suspected bots, and likely bots), e-cigarettes were more commonly mentioned than other tobacco products. Conversation about e-cigarettes was followed by conversation about cigarettes, with conversation around other products showing more variation across the three groups. Below are in-depth comparisons of product conversation, organized by geography. As is the case in any qualitative analysis, not all media mentions of e-cigarettes and other tobacco products refer to a specific product (for example, some posts may refer to smoking in general, without saying which type of product they are referring to); therefore, the percentages below should be understood as the proportion of conversation among all coded posts. The same is true for the coded themes, presented in a subsequent section.

Nationally

At the national level, likely bots showed the largest proportion of conversation around e-cigarettes (55%), compared to that of humans (49%) and then that of suspected bots (44%). Across all three groups, cigarettes followed as the most commonly mentioned product after e-cigarettes, with likely bots showing the lowest proportion of conversation around cigarettes (26%), followed by humans (31%) and then suspected bots (35%). Cigars followed as the next most commonly discussed product, with likely bots showing the largest proportion of conversation (13%), followed by humans (12%) and then suspected bots (11%). Conversation

around hookah made up a larger proportion of discourse among suspected bots (9%), compared to that of humans (7%) and that of likely bots (2%). Conversation around other tobacco products (cigarillos, little cigars, smokeless tobacco like chew and snus) was minimal, at less than 2.5% across all three groups.

New Jersey

Conversation in New Jersey showed slightly different patterns from those observed nationally. Humans showed the largest proportion of conversation around e-cigarettes (42%), compared to that of suspected bots (30%) and then that of likely bots (25%). As with the national trend, cigarettes were the next most commonly discussed product. Humans and suspected bots showed almost the same proportions of conversation around cigarettes (30%), followed by likely bots (24%). Discussion around cigars made up a substantially larger proportion of conversation among likely bots (24%) than it did for humans (11%) and for suspected bots (15%). Interestingly, conversation around smokeless tobacco was common among likely bots (22%), with suspected bots showing a far lower proportion of conversation (9%) and humans showing less than 0.5%.

Conversation by Product

United States

	Humans	Suspected Bots	Likely Bots
	48.6%	44.0%	55.3%
	30.8%	34.9%	26.7%
	11.6%	10.7%	12.8%
	7.3%	8.9%	2.2%
			2.3%

New Jersey

	Humans	Suspected Bots	Likely Bots
	42.1%	30.2%	24.6%
	29.5%	29.8%	24.1.%
	10.6%	15.2%	23.5%
	16.3%	15.2%	5.2%
		8.6%	22.0%

Legend: ● E-cigarettes ● Cigarettes ● Cigars ● Hookah ● Little Cigars/ Cigarillos ● Smokeless
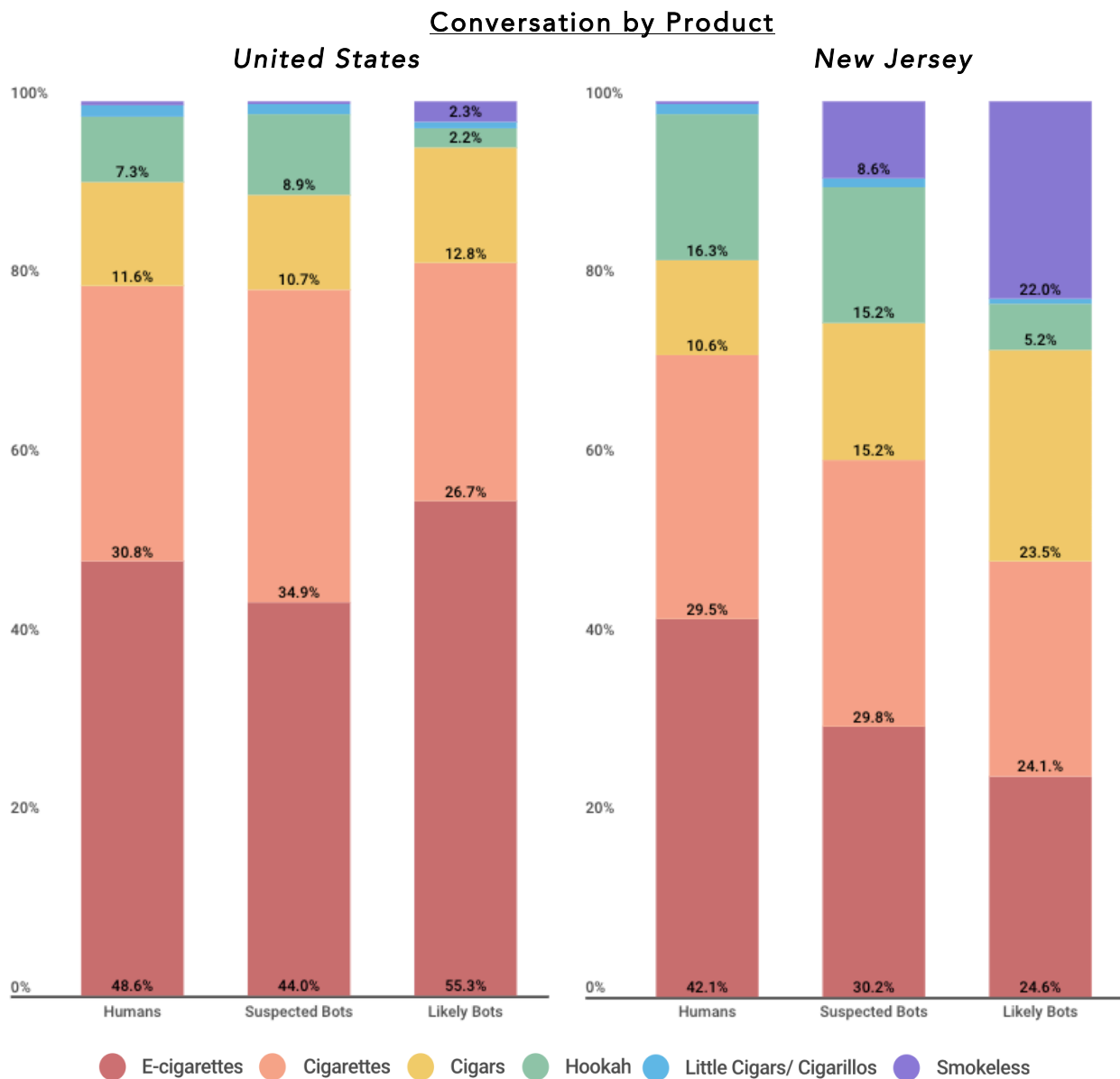
Analysis by Content Theme

The proportions of conversation dedicated to each theme showed interesting results, particularly within conversation by suspected bots. At the national level, conversation themes among suspected bots closely align with themes discussed by humans, while at the New Jersey level, conversation themes among suspected bots more closely mimic themes discussed by likely bots. Across both geographic areas, sales dominated conversation among likely bots, with the next most commonly discussed theme relating to quitting. Nationally,

likely bots showed more variations in conversation than likely bots in New Jersey, with the latter focusing over 80% of their conversation on sales and quitting. In both national and statewide conversation, humans dedicated about 15% of their conversation to sales, far lower than the percentage seen among likely bots. Humans tended to focus on quitting, health effects, and youth.
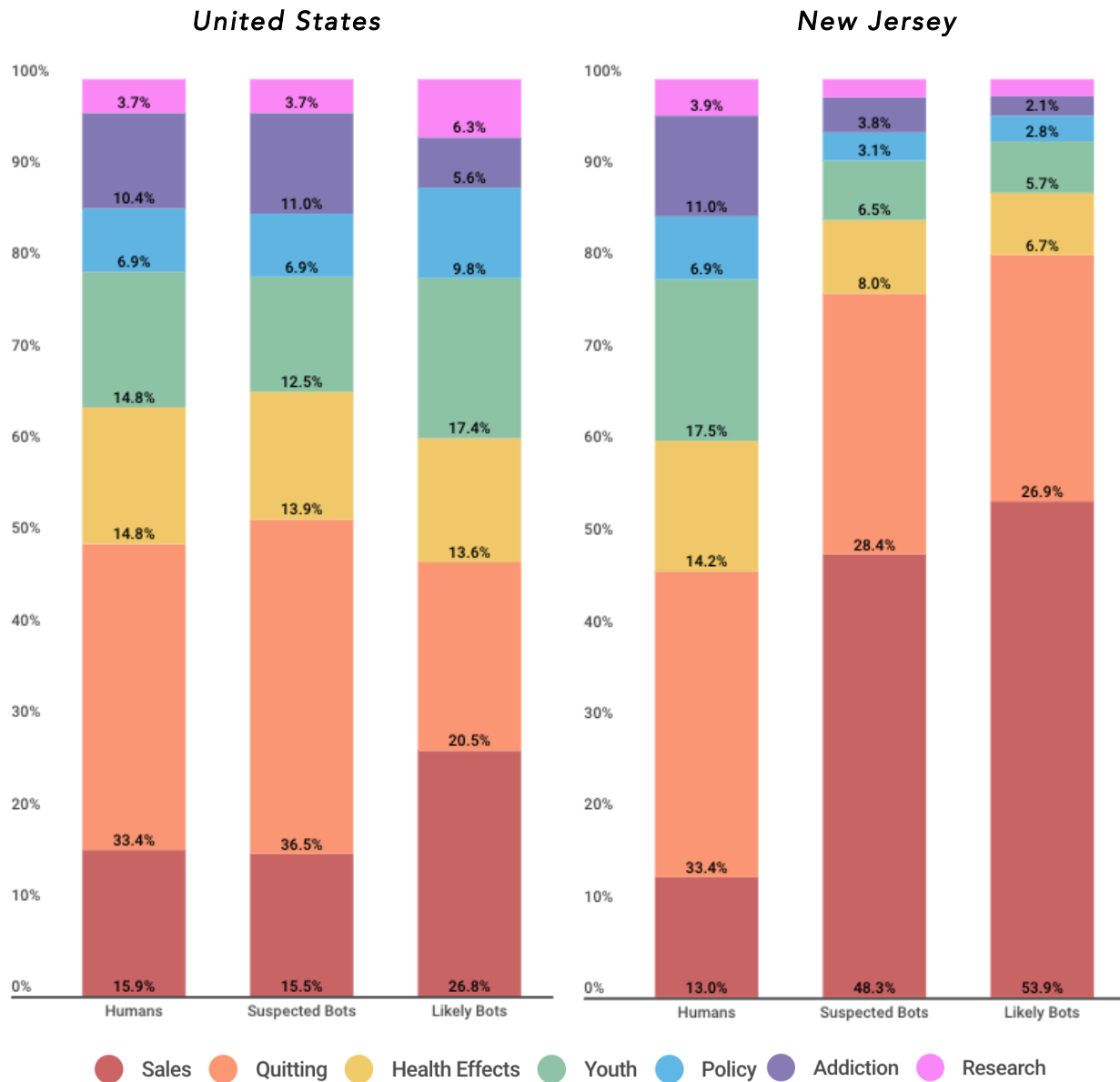
Nationally

At the national level, humans and suspected bots showed similarities in conversation themes, with nearly identical proportions of themes between the two. Conversation in these two groups tended to focus on quitting, with humans showing 33% of their conversation dedicated to this theme, compared to 37% among suspected bots and only 21% among likely bots. In contrast, likely bots showed a substantially higher proportion of their conversation to be around sales at 27%, compared to 16% among both humans and suspected bots. Conversation around the health effects of e-cigarettes and tobacco was nearly the same across all three groups (14%–15%), and likely bots showed a slightly higher proportion of conversation around youth (17%) than suspected bots (13%) and humans (15%). All other themes fell below 10% of conversation across the three groups, with likely bots showing a larger proportion of their conversation to be around policy, while humans and suspected bots showed a larger proportion of their conversation to be around addiction.

New Jersey

Conversation in New Jersey showed different trends. In New Jersey, suspected bots and likely bots showed similarities in conversation themes. Around half of their conversation was dedicated to discussing sales, with 54% among likely bots and 48% among suspected bots, compared to only 13% among humans. Conversation around quitting was more common among humans (33%) than among suspected bots (28%) and likely bots (27%). Humans also more often discussed topics related to youth (18%) than suspected bots (7%) and likely bots (6%). The proportion of conversation around health effects was also larger among humans (14%) than among suspected bots (8%) and likely bots (7%).

Conversation by Content Theme

United States

	Humans	Suspected Bots	Likely Bots
Research	3.7%	3.7%	6.3%
Addiction	10.4%	11.0%	5.6%
Policy	6.9%	6.9%	9.8%
Youth	14.8%	12.5%	17.4%
Health Effects	14.8%	13.9%	13.6%
Quitting	33.4%	36.5%	20.5%
Sales	15.9%	15.5%	26.8%

New Jersey

	Humans	Suspected Bots	Likely Bots
Research	3.9%		2.1%
Addiction	11.0%	3.8%	2.8%
Policy	6.9%	3.1%	5.7%
Youth	17.5%	6.5%	6.7%
Health Effects	14.2%	8.0%	26.9%
Quitting	33.4%	28.4%	53.9%
Sales	13.0%	48.3%	

Legend: ● Sales ● Quitting ● Health Effects ● Youth ● Policy ● Addiction ● Research

Commonly Used Hashtags

Hashtags are used within social media as a way of tagging conversations with a common topic or subject. When examined overall, they can be useful in garnering a quick visual snapshot of the types of discussions that occur within a specific topic. PGP analysts examined the 50 most commonly used hashtags within conversation among the three different groups. Across each

of the hashtag clouds below, analysts removed the terms #vape, #vaping, #vapefam, and #vapelife because they were consistently the top four hashtags. Removing these four allows for a deeper examination of hashtag trends throughout each of the groups.

Nationally

At the national level, the hashtag cloud from human conversation shows the top three hashtags as #worldnotobaccoday, #smoking, and #notobacco. Hashtags from humans tended to include both pro-tobacco cessation terms (#quitsmoking) and pro-vaping terms (#vapeon). In comparison, the top three hashtags among suspected bots were #ecig, #worldnotobaccoday, and #ejuice, respectively. Although this group included some anti-tobacco hashtags, most of the commonly used hashtags among suspected bots trended toward pro-vaping topics. Finally, among likely bots, the top three hashtags were #health, #smoking, and #ecig, respectively. These results appear to show a gradual transition from more anti-tobacco hashtags among humans to more pro-vaping messages among suspected bots and finally to positive, health-based hashtags, often in relation to the purportedly positive health benefits of using e-cigarettes, among likely bots. This examination of hashtags is important not only to understand the common types of topics discussed but also to track the frequency of types of conversation. By using the same hashtags as those used by bots, scientists can join the conversation with scientifically vetted information.

New Jersey

The most commonly used hashtags in New Jersey showed similar patterns. The top three hashtags used among humans were #smoking, #cigars, and #ecigs, respectively. Conversation among humans in New Jersey tended to be more pro-smoking and pro-vaping than conversation among humans at the national level. For example, the hashtag #worldnotobaccoday was the 10th most commonly used hashtag in New Jersey (compared to number one above). Among suspected bots in New Jersey, the top three hashtags were #smoking, #cigars, and #quitsmoking, respectively. Although the latter hashtag, #quitsmoking, appears to offer a positive pro-cessation message, it is often used in reference to using e-

cigarettes as a method of quitting smoking. Likely bots showed clear pro-e-cigarette and pro-cigar messages, with the top three hashtags being #vapenation, #cigars, and #cloudchaser, respectively. This group commonly uses sales-based hashtags, with #cloudchaser, #vapesale (number five most used), and #vapingcheap (number six most used) all predominantly referring to sales of e-cigarettes and related products online. Results show that anti-tobacco and anti-e-cigarette hashtags are relatively less common in New Jersey than across the nation, particularly among humans.

National Hashtag Clouds

Humans   *Suspected Bots*   *Likely Bots*

New Jersey Hashtag Clouds

Top Words Found in Commonly Shared News Stories

To further understand the types of conversation propagated by these three groups, PGP analysts then examined the most frequently used words within the most commonly shared

news stories and articles on Twitter. The word clouds below present the 100 words that were used most often in stories. The larger the word, the more often it was used.

Results from word clouds show some similarities between national and statewide conversations, with humans tending to focus on promoting policy-related news stories, suspected bots sharing content about sales, and likely bots sharing pro-vaping advocacy content that focuses on the positive aspects of e-cigarettes or that promotes the idea of quitting cigarettes and switching to e-cigarettes.

Nationally

At the national level, humans tended to share articles that focus on laws, legislative proposals, fines, or regulations related to e-cigarettes. The word "Ohio" is in reference to an Ohio lawmaker who wanted to ban smoking in vehicles with passengers younger than 6.[22] While suspected bots also shared the news article about the Ohio lawmaker, they also shared articles and sites related to sales, with the second most commonly shared article promoting e-cigarette sales on eBay. Likely bots showed vastly different patterns, with many of the words focusing on the experiential aspect of vaping, with words such as "sweet," "taste," "blueberry," and "experience" dominating their word cloud.

New Jersey

In New Jersey, as with the national trend, commonly shared articles among humans tended to focus on policies and laws related to e-cigarettes. Conversation also focused more heavily on the potentially negative aspects of vaping, with words such as "explosion" and "jaw" ranking higher on their word cloud; humans in New Jersey frequently shared an article referencing an instance in which a man died after his e-cigarette exploded in his mouth.[23] Suspected bots in

---

[22] Warwick, G. (2019). Ohio lawmaker wants to ban smoking in vehicles with passengers younger than 6. Retrieved June 23, 2019, from https://myfox28columbus.com/news/local/ohio-lawmakers-wants-to-prohibit-smoking-in-vehicles-with-passengers-younger-than-6

[23] Sarder, S. (2019). Fort Worth man dies after vape pen explodes at store, severs artery. Retrieved June 23, 2019, from https://www.dallasnews.com/news/fort-worth/2019/02/04/fort-worth-man-dies-after-vape-pen-explodes-store-severs-artery

New Jersey more commonly included words that relate to sales, with words such as "mod," "kit," and "40w" (in reference to e-cigarette voltage rates) ranking high in their word cloud. Suspected bots also tended to share links to eBay sales sites. Interestingly, the top article shared among likely bots was a story about a Hawaii lawmaker who proposed banning the sale of cigarettes to anyone under 100 years old.[24] This article is commonly shared in the context of pro-vaping advocates who post anti-cigarette messages. As a reflection of this story, the likely bots' word cloud included words such as "Hawaii," "proposed," and "ban," as well as sales-related terms, such as "sale" and "buy."
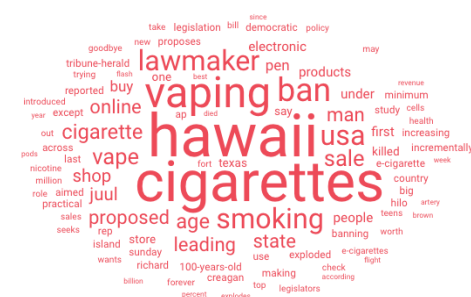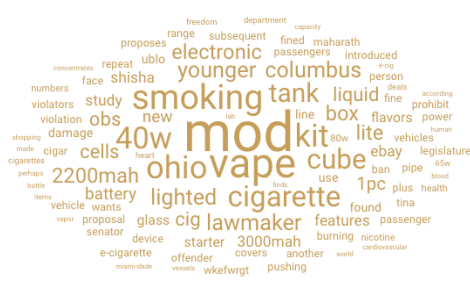
National Word Clouds

Humans   *Suspected Bots*

Likely Bots
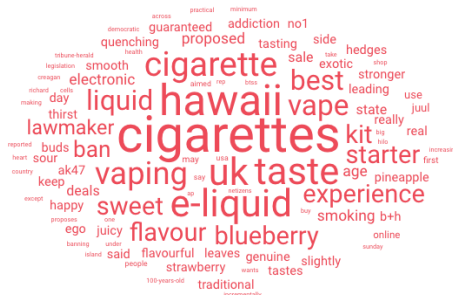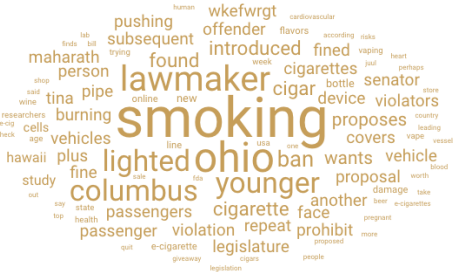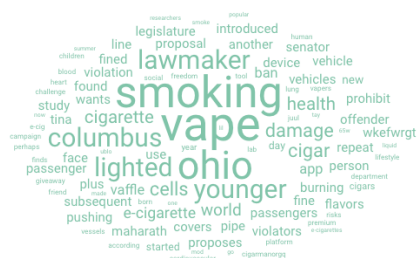
New Jersey Word Clouds

---

[24] Associated Press. (2019). Hawaii lawmaker proposes banning sale of cigarettes to anyone under 100-years-old. Retrieved June 23, 2019, from http://www.fox5dc.com/news/hawaii-lawmaker-proposes-banning-sale-of-cigarettes-to-anyone-under-100-years-old

Key Takeaways

Results from this analysis present never-before-seen information on the role that bots are currently playing in online conversation around e-cigarettes and tobacco products. Key takeaways from the findings above are summarized below:

- *Bots are playing a significant role in driving conversation at the national level as well as within New Jersey.*
    - Almost 80% of all conversations at the national level and almost 90% of conversations at the state level are driven by accounts that are either suspected or likely to be bots. This result corroborates what has been proposed in preliminary studies and suggests that the scope of the problem is even more alarming than initially believed. Future efforts should focus on informing the public of the vast role of bots in social media, particularly within conversation around e-cigarettes and tobacco products, and on promoting messaging about e-cigarettes that is based on scientifically vetted information.
- *Online conversation around e-cigarettes and tobacco products is complicated, and there is no one type of bot.*
    - Results showed an array of different types of malicious bots that are involved in conversation around e-cigarettes and tobacco. Bots can focus on creating groundswell movements based on misinformation, they can focus on sharing or retweeting information from other accounts, they can focus on simply promoting sales, or they can focus on a combination of these. While a recent Pew Research Center study showed that almost 70% of people in the United States are aware of the existence of bots, only 7% of them feel very confident that they can identify a bot on social media.[25] Given the enormity of bot influence on discourse and the array of different types, it is important to keep in

---

[25] Stocking, G., & Sumida, N. (2018). Social media bots draw public's attention and concern. Pew Research Center, Washington, DC.

mind the complexity of bot conversation and that most internet users are likely unable to determine the difference between bots and humans.

- *Discussion around e-cigarette sales is a substantial piece of conversation, particularly among accounts that are likely bots.*
  - o Research has shown that the e-cigarette market continues to undergo an explosion in sales as companies innovate and expand. Online sales make up a substantial portion of the overall market for both adults and adolescents, with estimates that over one-third of teens who use e-cigarettes purchase them either in a store or online.[26][27] Despite the fact that each state has its own laws restricting sales of e-cigarettes to minors, two recent studies have shown that between 60% and 98% of e-cigarette sales transactions attempted by a minor did not employ reliable age verification techniques before purchase.[28,29] Results from this report suggest that it may be useful to adapt the framing and content of sales messages for counter-message development as a way of resonating with the audience and providing accurate information around vaping. It may also be useful to utilize common hashtags that are found within conversation around vaping sales as a way of joining the conversation with scientifically vetted messaging.

- *Bots are actively promoting smoking cessation in exchange for vaping.*
  - o One topic of conversation commonly observed among bots referenced the importance of quitting smoking. While on the surface this appears to be a positive finding, a deeper examination discovers a majority of the conversation

[26] Pepper, J. K., Coats, E. M., Nonnemaker, J. M., & Loomis, B. R. (2018). How do adolescents get their e-cigarettes and other electronic vaping devices?. American Journal of Health Promotion, *33*(3), 420–429. [0890117118790366].

[27] VynZ Research. (2019). Global E-cigarette market is Set to Reach USD 53.4 billion by 2024, Observing a CAGR of 21.4% during 2019–2024: VynZ Research. *Global News Wire.* Retrieved September 16, 2019, from https://www.globenewswire.com/news-release/2019/05/20/1827594/0/en/Global-E-cigarette-market-is-Set-to-Reach-USD-53-4-billion-by-2024-Observing-a-CAGR-of-21-4-during-2019-2024-VynZ-Research.html

[28] Williams, R. S., Derrick, J., Liebman, A. K., LaFleur, K., & Ribisl, K. M. (2018). Content analysis of age verification, purchase and delivery methods of internet e-cigarette vendors, 2013 and 2014. *Tobacco Control, 27*(3), 287–293.

[29] Williams, R. S., Derrick, J., & Ribisl, K. M. (2015). Electronic cigarette sales to minors via the internet. *JAMA Pediatrics, 169*(3), e1563.

around quitting smoking exists to promote vaping. It was common to see bots embed anti-tobacco sentiment within pro-vaping messaging. This is important to know for messaging that counters the information promoted by bots. At face value, bot messaging can appear to be anti–tobacco industry, anti-smoking, and pro-health. Public health messaging that highlights the transgressions of the tobacco industry may actually be the same content that people are seeing from bots. This nuance in content is important to understand, as bots can appear to promote messages aligned with public health when in reality they are pushing a pro-vaping agenda.

- *Pro-vaping advocacy is common and current.*
    - Bots appear to be active in rallying against policies that would negatively impact e-cigarettes, including policies related to sales, access, and production. Bots tend to keep up with current events related to policy changes and legislation—both nationally and internationally—and actively promote humans becoming involved in efforts to discourage changes in policy. It is therefore important that messaging to counter this bot advocacy be just as current as the bots themselves. There is no organization monitoring bot involvement in e-cigarette advocacy campaigns, which is particularly alarming given that bots have recently been proven to have a tremendous impact on other important areas of national interest, including policy and politics.[30,31,32] While bots' influence on politics is just beginning to be understood and addressed, their power and potential ability to sway conversation around e-cigarettes appears to be going relatively unchecked.

- *Misinformation is a common—and worrying—portion of conversation.*

---

[30] Woolley, S. C. (2016). Automating power: Social bot interference in global politics. *First Monday*, *21*(4).

[31] Hern, A. (2017). Facebook and Twitter are being used to manipulate public opinion–report. *The Guardian.* Retrieved June 26, 2019, from https://www.theguardian.com/technology/2017/jun/19/social-media-proganda-manipulating-public-opinion-bots-accounts-facebook-twitter

[32] Stella, M., Ferrara, E., & De Domenico, M. (2018). Bots increase exposure to negative and inflammatory content in online social systems. *Proceedings of the National Academy of Sciences, 115*(49), 12435–12440.

o Several bot accounts presented above focus on promoting misinformation regarding e-cigarettes. These accounts expressly accuse scientists of hiding the truth and actively refute and discredit information spread by public health organizations, such as the Centers for Disease Control and Prevention and the Food and Drug Administration, among other key players in public health. With a lack of vetted scientific information on e-cigarettes, there is a risk that humans are seeing this misinformation without facts with which to compare it. These accounts underscore the vital importance of continued monitoring to alert public health authorities and practitioners to the common types of misinformation driven by bots so that they are able to counter that messaging with timely, relevant content of their own.