



Selected Papers of #AoIR2019:
The 20th Annual Conference of the
Association of Internet Researchers
Brisbane, Australia / 2-5 October 2019

BLIND SPOTS OF INFORMATION OPERATIONS: OF MICRO PROPAGANDA, ALGORITHM GAMING & HOW TO PROFIT FROM IT

Panel Introduction

The same technologies that once promised to strengthen democracy increasingly have become accused of undermining political processes. The Internet, and specifically social media—once celebrated as technologies of liberation—have recently shown themselves to be vulnerable hosts to various forms of influence campaigns and manipulation. Around the world, a wide range of actors exploit social media platforms to distort public discourse and undermine trust. Malicious foreign influence campaigns armed with “fake news” and automation, hyperpartisan alternative media outlets propagating conspiracy and hate, and extremist voices at the fringe of the spectrum all have found a fertile breeding ground in social media networks.

Already, scholars have argued that the spread of malicious information campaigns and computational propaganda sustained by social media algorithms can corrode democratic discourse and disrupt digital public spheres (Persily, 2017; Tucker, Theocharis, Roberts, & Barberá, 2017; Woolley & Howard, 2016). A growing body of scholarship, has become concerned with tracking the spread disinformation and automated messages during sensitive moments of public life (Ferrara et al., 2016; Vosoughi, Roy & Aral, 2018) the relationship between social media algorithms and the viral dissemination of various kinds of problematic information (Bradshaw & Howard, 2018; Silverman, 2015; Wu, 2017); the different kinds of nefarious actors behind political influence campaigns (Marwick & Lewis, 2017; Faris et al., 2017); and possible countermeasures and solutions (Ash, Gorwa, & Metaxa, 2019; Bradshaw, Neudert & Howard, 2018).

But techniques designed to manipulate public opinion and undermine information ecosystems are rapidly evolving while academic research lags behind technological innovation and strategic expertise. As a new and more sophisticated generation of information operations is fast to mature, the papers in this panel shed light on some of the blind spots of scholarly inquiry making visible new thematic strategies, technical infrastructures, and both political and economic incentives.

The first two papers examine the progression from general political propaganda geared towards influencing elections to highly issue-specific micro propaganda. The first paper

presents an analysis of anti-Semitic disinformation campaigns and political harassment during the 2018 US midterms on Twitter and offers rich evidence from interviews with Jewish American opinion leaders about the impact of these campaigns. Drawing on data from Twitter's Election Integrity Initiative, the second paper examines the gender dimensions of foreign influence operations and how hostile state actors frame and discuss gender identity & politics. The third paper presents an analysis of search engine optimization strategies that extremist YouTubers use in an attempt to game the algorithm and increase their visibility in the network. The fourth paper investigates the relationship between partisan bias associated with Google Search results and the success of political candidates associated with the search queries during elections and finds that partisan search media is a predictor for election outcomes. The fifth paper examines the emergence of a global political economy for manipulation and offers a grounded typology of the vendors, marketplaces, services, and products that are designed to turn a profit from swaying public opinion.

Together, the papers on the panel present methodological research into widely unexplored phenomena having to do with information operations, including issue-specific influence campaigns, search engine hacking, and profit-driven manipulation. While journalistic and expert inquiries have reported on several phenomenological instances discussed in these papers, the debate so far widely lacks scholarly attention. Rooted in political communication literature on disinformation, "computational propaganda" and hyperpartisan media, the papers offer data-driven research into both technological systems and the actors that seek to manipulate them. Thus, this panel analyzes through an exploratory lens illuminating technological, thematic and strategic blind spots of state-of-the-art information operations.

References

- Ash, T. G., Gorwa, R., & Metaxa, D. (2019). *GLASNOST! Nine ways Facebook can make itself a better forum for free speech and democracy*. Oxford. Retrieved from https://pacscenter.stanford.edu/wp-content/uploads/2019/01/Garton_Ash_et_al_Facebook_report_FINAL_0.pdf
- Bradshaw, S., & Howard, P. N. (2018). Why does Junk News Spread So Quickly Across Social Media? Algorithms, Advertising and Exposure in Public Life. *Knight Foundation Working Paper*. Retrieved from https://kf-site-production.s3.amazonaws.com/media_elements/files/000/000/142/original/Topos_KF_White-Paper_Howard_V1_ado.pdf
- Bradshaw, S., Neudert, L. M., & Howard, P. N. (2019). *Government Responses To Malicious Uses of Social Media* (Working Paper 2019.2). Oxford, UK: Oxford University, Project on Computational Propaganda.
- Faris, R. M., 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 for Internet & Society Research Paper. Retrieved from https://dash.harvard.edu/bitstream/handle/1/33759251/2017-08_electionReport_0.pdf

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The Rise of Social Bots. *Commun. ACM*, 59 (7), 96–104. <https://doi.org/10.1145/2818717>

Marwick, A., & Lewis, R. (2017). *Media Manipulation and Disinformation Online*. Data & Society Research Institute. Retrieved from https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

Persily, N. (2017). The 2016 U.S. Election: Can Democracy Survive the Internet? *Journal of Democracy*, 28(2), 63–76.

Silverman, C. (2015). *Lies, Damn lies, and viral content. How news websites spread (and debunk) online rumors, unverified claims, and misinformation* (p. 164). Tow Center for Digital Journalism.

Tucker, J. A., Theoharis, Y., Roberts, M. E., & Barberá, P. (2017). From liberation to turmoil: social media and democracy. *Journal of Democracy*, 28(4), 46–59.

Woolley, S. C., & Howard, P. N. (2016). Political Communication, Computational Propaganda, and Autonomous Agents. *International Journal of Communication*, 10, 4882–4890.

Wu, T. (2017). *The Attention Merchants: The Epic Struggle to Get Inside Our Heads*. Atlantic Books.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>

UNDERSTANDING THE HUMAN IMPACT OF COMPUTATIONAL PROPAGANDA: ANTI-SEMITISM AND THE 2018 US MIDTERM ELECTIONS

Samuel Woolley
Digital Intelligence Lab, Institute for the Future

Katie Joseff
Digital Intelligence Lab, Institute for the Future

Introduction

Jeff, a Jewish American reporter at a major U.S. news outlet, first experienced anti-Semitic political attacks online during the 2016 U.S. presidential election. When his stories gained traction—or featured details on topics such as white nationalism, Donald Trump, or libertarianism—he would be sent photoshopped images of his face in a gas chamber or messages threatening violence. Now, this harassment has become common. “It has become something that I expect to happen,” he said, “I don’t even think about it anymore.” For Jeff, the 2016 election was a major impetus that transformed the online sphere into a hotbed of hate speech and harassment.

The number of Jewish people living in the United States is estimated to be between 4.2 million and 12 million, the wide range due to religious versus ethnic distinctions (Steinhardt Social Research Institute, 2016; DellaPergola, 2017). For many of these individuals—especially those in the public eye—social media platforms have become inhospitable for both general communication and as forums for discussing public life. This report explores the ways in which online propaganda, harassment, and political manipulation have impacted Jewish Americans during and after the 2018 U.S. midterm elections. In the course of our research, interview subjects have described a marked rise in the number of online attacks they have experienced. Correspondingly, our data analyses suggests that tools like social media bots and tactics, including doxxing, disinformation, and politically-motivated threats, have been used to target Jewish Americans.

Literature Review

There has been an undisputed rise in white supremacist activities and overt anti-Semitism following the 2016 U.S. election (ADL, 2018b). From 2016 to 2017, the number of established neo-Nazi groups increased from 99 to 121 (SPLC, 2018); twice as many hate-motivated murders were committed by white supremacists (Baynes, 2018); and there was a 258% increase in the number of white supremacist propaganda incidents on college campuses (ADL, 2018a). While not all white supremacist groups consider themselves anti-Semitic, anti-Semitism is often a core tenet of white supremacy and, by extension, white nationalism and neo-Nazism (Ferber, 1999). As such, it comes as little surprise that 1,986 anti-Semitic incidents—harassment, vandalism, and assault—occurred in 2017 (ADL, 2018b). The 57% increase in such events was the largest escalation in a single year since the Anti-Defamation League (ADL) first began recording incidents in 1979.

A staggering expansion of online harassment coincided with, and arguably fomented, the increase in offline anti-Semitism. Fringe Internet communities, such as 4chan,

8chan, and Gab allowed for the propagation of such ideas, which quickly spread to Twitter, Reddit, and other mainstream online communities (Marwick & Lewis, 2017). An analysis of over 100 million posts on Gab and 4chan's Politically Incorrect message board (/pol/) found that, between July 2016 and January 2018, the use of the terms "Jew" and "kike," a derogatory term for Jewish people, more than doubled on /pol/ and dramatically increased on Gab (Finkelstein et al., 2018). Spikes also occurred in the use of both terms following President Trump's inauguration and the 2017 Unite the Right Rally in Charlottesville (ADL, 2017).

Methodology

Expanding upon our preliminary report published by the ADL in 2018 (Woolley & Joseff, 2018), the intention of this report is to better understand the impact of online political harassment and anti-Semitic disinformation during and after the 2018 U.S. midterm elections. We conducted both interviews and analyses of tweets in order to understand the scope of the issue on a national scale and the repercussions faced on the individual level.

We interviewed eighteen Jewish Americans who are involved in American politics as elected officials, policy makers, activists, journalists, and consultants. While somewhat limited in number, the interviewees were diverse in political affiliation, Jewish movements membership, age, and race. For our data analysis, we selected for specific hashtags (e.g. #MAGA and #VoteBlue) and collected 7,512,594 tweets related to U.S. politics. Using Tweepy, the tweets were gathered in groups between August 31, 2018 and September 17, 2018. The hashtags were categorized by political leaning: conservative, liberal, extremist, and neutral (e.g. "#vote"). The hashtags were purposively gathered using markers from previous and ongoing research on Twitter conversations (Kollanyi et al., 2016; Woolley & Guilbeault, 2017). We worked to be non-partisan in our selection of hashtags and analyses of data in order to produce the most objective results possible; although, we accept the impossibility of true positivism in social scientific research.

Tweets were then filtered based upon whether or not the text of the tweet contained a series of terms related to Judaism and/or anti-Semitism. Instances of term use in hashtags, usernames, and shared links were not included. The terms were categorized as: derogatory, lean derogatory, context dependent, lean context dependent, neutral (e.g. Jew, Orthodox, Israeli), and other, which consisted of derogatory terms historically used by Jews to describe other ethnicities, non-Jewish individuals, and to criticize other Jews (e.g. kushi, kapo). The accounts that tweeted five or more derogatory or lean derogatory terms during the time period were then run through Botometer to ascertain whether or not they were automated.

Implications

The most startling trend is the transmutation of online hate speech into real world violence. Many of the graphic memes that were used to target our interviewees were created on 4chan and spread through targeted Twitter campaigns initiated by white supremacist leaders. Often these campaigns involved doxxing—the release of the target's private information—which increases the likelihood of offline violence. The immense stress caused by these attacks is difficult to quantify.

References

- ADL. (2017, August 7). "Unite the Right" Rally Could Be Largest White Supremacist Gathering in a Decade. Retrieved from <https://www.adl.org/blog/unite-the-right-rally-could-be-largest-white-supremacist-gathering-in-a-decade>
- ADL. (2018a). ADL Finds Alarming Increase in White Supremacist Propaganda on College Campuses Across U.S. Retrieved from <https://www.adl.org/news/press-releases/adl-finds-alarming-increase-in-white-supremacist-propaganda-on-college-campuses>
- ADL. (2018b). Audit of Anti-Semitic Incidents: Year in Review 2017. Retrieved from <https://www.adl.org/media/11174/download>.
- Baynes, C. (2018). Murders by white supremacists in America more than doubled last year. Retrieved from <https://www.independent.co.uk/news/world/americas/white-supremacist-murders-us-figures-double-2017-racist-hate-crime-las-vegas-shooting-extremist-a8165416.html>.
- DellaPergola, S. (2017). World Jewish Population, 2016. In American Jewish Year Book 2016: The Annual Record of North American Jewish Communities (Vol. 116, pp. 253–332).
- Ferber, A. L. (1999). *White Man Falling: Race, Gender, and White Supremacy*. Lanham, MD: Rowman & Littlefield Publishers.
- Finkelstein, J., Zannettou, S., Bradlyn, B., & Blackburn, J. (2018). A Quantitative Approach to Understanding Online Antisemitism. arXiv preprint arXiv:1809.01644.
- Marwick, A., & Lewis, B. (2017). *Media Manipulation and Disinformation Online*. New York, NY: Data & Society Research Institute.
- SPLC. (2018). The Year in Hate: Trump buoyed white supremacists in 2017, sparking backlash among black nationalist groups. Retrieved from <https://www.splcenter.org/news/2018/02/21/year-hate-trump-buoyed-white-supremacists-2017-sparking-backlash-among-black-nationalist>
- Steinhardt Social Research Institute. (2016). American Jewish Population Project. Waltham, MA: Cohen Center for Modern Jewish Studies, Brandeis University. Retrieved from <http://ajpp.brandeis.edu/aboutestimates.php>
- Woolley, S., & Guilbeault, D. (2017). Computational Propaganda in the United States of America: Manufacturing Consensus Online (Computational Propaganda Working Paper Series No. 2017.5) (p. 37). Oxford, United Kingdom: Oxford Internet Institute, University of Oxford.
- Woolley, S., & Joseff, K. (2018). Computational Propaganda, Jewish-Americans and the 2018 Midterms: The Amplification of Anti-Semitic Harassment Online. New York, NY:

Anti-Defamation League. Retrieved from
<https://www.adl.org/resources/reports/computational-propaganda-jewish-americans-and-the-2018-midterms-the-amplification>

THE GENDER DIMENSIONS OF FOREIGN INFLUENCE OPERATIONS

Samantha Bradshaw
Oxford Internet Institute, University of Oxford

Introduction

Malicious state actors are increasingly leveraging social media as a proxy for political power. During elections and other important political events, coordinated disinformation campaigns have been used to manufacture consensus, automate suppression, and undermine trust in media, institutions and democracy (Bradshaw & Howard, 2017, 2018a). One aspect of foreign influence operations that has received little scholarly and public attention has been on the gender dimensions of these campaigns. It is widely recognized that female politicians, journalists, bloggers, and activists—especially those of colour or diverse gender identity—are frequently targeted with intimidation, harassment, threats, and hate speech in online spaces (Amnesty International, 2018). Yet little research has examined how malicious state actors are using trolling tactics and disinformation to perpetuate sexism and misogyny in order to suppress the political speech of vulnerable groups, heighten cultural tensions, and further divisions within and across communities. Drawing on data from Twitter's *Election Integrity Initiative*, this paper explores how gender identity & politics are framed and discussed by foreign state actors. By performing an analysis of prominent keywords and hashtags related to gender identity and gender-related political movements, this paper hopes to provide insight into the gender dimensions of foreign influence operations.

Literature Review

Foreign influence operations on social media have emerged as a critical concern of the 21st century. Since 2016, many researchers and journalists have uncovered foreign influence operations taking place on social media during critical moments of public life. These studies have mainly focused on the Internet Research Agency's (IRA) disinformation campaigns, situating modern strategy with historical tenants (Maréchal, 2017), describing the Russian playbook for information warfare (Polyakova, Laurelle, Meister, & Barnett, 2016), or analysing their broader campaigns in the Baltics, for example (Helmus, 2018). This quickly growing topic of study has also expanded to look at disinformation campaigns taking place across a wide-range of platforms, country-contexts, and issue-areas including race, religion, and other social justice issues (Marwick & Lewis, 2017; Woolley, 2016; Woolley & Howard, 2018).

Although foreign influence operations have always been a part of political and military strategy, the unique characteristics of the Internet and social media coupled with advances in technology are posed to give rise to a new generation of tools and techniques that expand the scope and scale of foreign influence operations (Hwang & Rosen, 2017). The growing study of "computational propaganda" (Howard & Woolley, 2016) highlights how both user capacity and technological affordances both enhance and constrain the spread of disinformation about politics (Bradshaw & Howard, 2018b). But the "social shaping" (Mackenzie & Wajcman, 1985) of social networking technologies also have implications for the spread of hate speech and misogyny against women and those of diverse gender identities (Banet-Weiser & Miltner, 2016; Mantilla, 2013). As hostile state actors continue to look for pressure points within society, it is important to understand

how other groups or communities such as women or individuals with diverse gender identities, might be affected by foreign influence operations that are sustained by the affordances of social media platforms. This research paper asks three questions: (1) What are the gendered elements of foreign influence operations? (2) How do the technical affordances, structures and policies of online platforms, as well as existing legal frameworks, enhance or constrain the gendered elements of foreign interference? and (3) How best can policymakers and social media platforms respond to these challenges?

Methodology

Building on the growing corpus of research about foreign influence operations on social media, this paper explores the gender dimensions of these campaigns. This study focuses specifically on influence operations that have taken place on Twitter. Twitter is an important social media platform for political news and discussion (Newman, Fletcher, Kalogeropoulos, Levy, & Nielsen, 2018). As a result, several hostile state actors have exploited Twitter to conduct online influence operations. Although these activities occur on all major social media platforms, Twitter was chosen as the platform to analyse because the company has released the most comprehensive dataset about foreign influence operations to date. *Twitter's Election Integrity Initiative* (Twitter, 2018) provides researchers with publicly available data about accounts and related content associated with potential influence operations that have taken place on the platform. This dataset includes information about accounts themselves (such as username, number of followers, and profile information), tweets (such as public posts and hashtags), and shared media (including pictures, video's and URLs). At the time of writing, Twitter has uploaded seven datasets about potential influence operations from four different countries: Russia, Iran, Venezuela and Bangladesh. These datasets form the foundation of my study.

To develop a better understanding on how gender identity & politics are used in discussions on Twitter by foreign state actors, this paper conducts a qualitative hashtag and keyword analysis to identify relevant discussions by hostile state actors. By purposively selecting keywords related to gender, I searched the database for relevant conversations. From there, I used a snowball sample to further identify collocated hashtags about gender identity and politics. Following this topic-based sampling strategy, I collected approximately 1.3 million Tweets from Twitter's Integrity Initiative dataset. Five prominent cross-cutting themes were identified: (1) gender, race and crime; (2) gender and Islam; (3) LGBT rights; (4) the feminist movement; (5) white male supremacy. At the time of writing, I am preparing a sentiment analysis to analyze the valence of these Tweets to see how gender is performed, framed and discussed by foreign state actors engaged in influence operations on Twitter. Sentiment analysis is a growing area of Natural Language Processing that determines whether a piece of text is positive, negative or neutral. In particular, I will adopt a "lexicon-based approach" to identify the sentiment of Tweets about gender politics. The results of these findings will be discussed in the context of policy responses to foreign influence operations, and enrich the academic discussion currently taking place in the field of political communication.

References

Amnesty International. (2018). *Toxic Twitter - A Toxic Place for Women*. Retrieved from <https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-1/>

Banet-Weiser, S., & Miltner, K. M. (2016). #MasculinitySoFragile: culture, structure, and networked misogyny. *Feminist Media Studies*, 16(1), 171–174.

Bradshaw, S., & Howard, P. N. (2017). *Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation* (Working Paper 2017.12 No. 2017.12) (p. 37). Oxford, England: Project on Computational Propaganda, Oxford Internet Institute, Oxford University.

Bradshaw, S., & Howard, P. N. (2018a). *Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation* (Working Paper 2018.1 No. 2018.1) (p. 26). Oxford, England: Project on Computational Propaganda, Oxford Internet Institute, Oxford University.

Bradshaw, S., & Howard, P. N. (2018b). Why does Junk News Spread So Quickly Across Social Media? Algorithms, Advertising and Exposure in Public Life. *Knight Foundation Working Paper*.

Helmus, T. C. (2018). *Russian social media influence: understanding Russian propaganda in Eastern Europe*. Santa Monica, Calif: RAND Corporation.
Howard, P., & Woolley, S. (2016). Political Communication, Computational Propaganda, and Autonomous Agents. *International Journal of Communication*, 10(Special Issue), 20.

Hwang, T., & Rosen, L. (2017). Harder, Better, Faster, Stronger: International Law and the Future of Online PsyOps. *Computational Propaganda Project Working Paper, Paper No.*

Mackenzie, D., & Wajcman, J. (1985). *The Social Shaping of Technology*. Milton Keynes: Open University Press.

Mantilla, K. (2013). Gendertrolling: Misogyny Adapts to New Media. *Feminist Studies*, 39(2), 563–570.

Maréchal, N. (2017). Networked Authoritarianism and the Geopolitics of Information: Understanding Russian Internet Policy. *Media and Communication*, 5(1).

Marwick, A., & Lewis, R. (2017). *Media Manipulation and Disinformation Online* (pp. 1–106). Data and Society.

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A. L., & Nielsen, R. K. (2018). Digital News Report 2018, 144.

Polyakova, A., Laurelle, M., Meister, S., & Barnett, N. (2016). *The Kremlin's Trojan Horses*. Atlantic Council.

Twitter. (2018). Elections integrity. Retrieved from https://about.twitter.com/en_us/values/elections-integrity.html

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

Woolley, S., & Howard, P. (2018). *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*. Oxford: Oxford University Press.

SEARCH ENGINE OPTIMIZATION ON POLITICAL YOUTUBE: SOCIO-TECHNICALLY SITUATED DISINFORMATION AND PROPAGANDA

Rebecca Lewis
Stanford University

Leon Yin
New York University

In recent years, search engines have become invaluable resources for people seeking information. As described by information studies scholars Deirdre Mulligan and Daniel Griffin, search engines no longer only shape public understandings of digital content; they also “shape [our] public understanding of the world” at large (2018: 557). A range of quantitative and qualitative studies over the past decade have shown that people inherently trust the information that appears to them in online search results (Daniels, 2009; Purcell, Brenner, & Rainie, 2012; Tripodi, 2018). In this paper, we argue that despite this trust, search engines are themselves often instruments of propaganda, disinformation, and manipulation.

Specifically, we argue that political influencers and alternative news sources on YouTube are adopting the tactics of search engine optimization in ways that intentionally and unintentionally amplify misleading and problematic information – or place it within misleading or problematic contexts. Our analysis centers on the content tags that a range of political YouTube creators place on their videos to optimize their placement in search results, as well as the content of the videos that are labeled with these tags. Combining quantitative and qualitative analysis, we develop a typology for these tags to specify how they can result in differing kinds of misleading information.

While journalistic and academic accounts of search engines have largely focused on Google and Bing, these accounts neglect to account for the fact that YouTube is a crucial source for information-seeking, particularly among young internet users. According to a 2018 Pew Research Report on social media use in 2018, YouTube is used by nearly three-quarters of American adults and 94% of 18-to-24-year-olds (Smith & Anderson, 2018). About a third of those users get their news there, making it the second most common social media site for news, behind Facebook (Shearer & Gottfried, 2017). Most importantly for the purposes of this paper, YouTube is the second-most-popular search engine on the internet, behind online its parent company, Google (Richards, 2018).

More recently, analyses have begun to treat YouTube as a serious source of news and political content (Lewis, 2018; Tripodi, 2018; Tufekci, 2018). In particular, these analyses have focused on the radicalizing potential of right-wing political content on the platform. Tufekci’s work has highlighted the power of the YouTube algorithm to draw users into more extremist content, while Lewis and Tripodi have focused on the power of YouTube influencers as alternative news sources that can lead their audiences to extremist content.

Rather than focusing on recommendation algorithms or video content as pathways of radicalization, we focus on YouTubers' tactics to influence users' information-gathering at a first point of entry: YouTube search results. Thus, we also build on the growing body of literature that argues that as advertising-driven companies, search engines are not neutral distributors of information (Mulligan & Griffin, 2018; Noble, 2018). Specifically, as this literature has noted, a range of actors have learned to adopt "black hat" search engine optimization tactics that can exploit search results with low content results or mask the original sourcing of highly ranked results.

Our analysis builds on, and draws together, these two strands of literature by interrogating the search engine optimization strategies of political YouTubers and channels. We were driven by the following research questions: what specific SEO strategies do political YouTubers adopt to appear highly on YouTube search results? What are the possible implications for disinformation, propaganda, and other harmful information on the platform?

To answer these questions, we developed a population of analysis by identifying seed English-language YouTuber accounts on both the political right and the left; we then used a snowball method to build out the larger sample. We extracted search tags from video metadata retrieved from the YouTube data API. These tags are explicitly input by users uploading videos, and they help the YouTube search algorithm rank results by relevance on related search terms. Although these tags are not visible to content viewers on the site, they are available from the API. For this reason, for any given user, we can count the unique instances of search tags across every video they have uploaded.

For videos with specific tags, we also performed a qualitative analysis of the video content itself. We watched a wide range of videos within the population, and, in certain cases, we manually downloaded and performed qualitative coding on YouTube's automated video transcripts. This allowed us to combine quantitative metrics from video tags with qualitative insights from videos themselves, thus giving us better insight into the work accomplished by each tag.

We developed a typology of search engine optimization tactics employed by these YouTubers. First, we observed evidence of "*issue hacking*," a concept previously explored by Ried, Matamoros-Fernández, and Coromina (2018), in which YouTubers strategically tag their content with hot-button issues or newsworthy topics to attract more viewers. Second, we identified attempts at *influencer amplification*, in which users tag their content with specific influential people that may be able to further amplify their content – or who may be frequent subjects of searches. Finally, we identified keyword tags that operate specifically as *dog whistles and gateways*, in which content becomes matched with terms that specifically mask extremist messaging.

In all cases, we argue that the disinformation and propaganda within certain videos has the potential to become far more potent when successfully placed within certain search contexts. Overall, our findings indicate that digital content creators understand the importance of the positioning of information within certain socio-technical contexts. As a result, it is essential for researchers to move beyond understandings of digital

propaganda and disinformation that focus solely on the content of messages. Disinformation is not only a matter of content that needs to be fact-checked, but is also a result of complex relationships between users and the technical systems they use.

References

Daniels, J. (2009). *Cyber Racism: White Supremacy Online and the New Attack on Civil Rights*. Lanham, MD: Rowman & Littlefield Publishers.

Lewis, R. (2018). *Alternative Influence: Broadcasting the Reactionary Right on YouTube* (White paper). Data & Society Research Institute.

Mulligan, D. K., & Griffin, D. S. (2018). Rescripting Search to Respect the Right to Truth. *Georgetown Law Technology Review*, 2(2), 557–584.

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York, N.Y.: NYU Press.

Purcell, K., Brenner, J., & Rainie, L. (2012). *Search Engine Use 2012*. Pew Research Center. Retrieved from <http://www.pewinternet.org/2012/03/09/search-engine-use-2012-2/>

Richards, L. (2018, June 27). Video and search: YouTube, Google, the alternatives and the future. Retrieved February 24, 2019, from <https://searchenginewatch.com/2018/06/27/video-and-search-youtube-google-the-alternatives-and-the-future/>

Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018). From ranking algorithms to 'ranking cultures': Investigating the modulation of visibility in YouTube search results. *Convergence*, 24(1), 50–68. <https://doi.org/10.1177/1354856517736982>

Shearer, E., & Gottfried, J. (2017, September 7). News Use Across Social Media Platforms 2017. Retrieved March 11, 2018, from <http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/>

Smith, A., & Anderson, M. (2018). *Social Media Use in 2018*. Pew Research Center.

Tripodi, F. (2018). *Alternative Facts, Alternative News: Conservatives search for Truth through Scriptural Inference*. Data & Society Research Institute.

Tufekci, Z. (2018, March 10). YouTube, the Great Radicalizer. *The New York Times*. Retrieved from <https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html>

INVESTIGATING PARTISANSHIP IN U.S. POLITICAL WEB SEARCH

Danaë Metaxa
Stanford University

Abstract

Concern about algorithmically-curated content and its impact on democracy is reaching a fever pitch worldwide. But relative to the role of social media in electoral processes, the role of search results has received less attention. This work analyzes news media in web search results pages in the context of political partisanship. Our empirical analyses use URLs scraped from Google search queries for political candidates in the 2018 U.S. elections. We use these data to empirically study trends theorized in political science, in particular finding that candidates with less partisan search media are more likely to win elections.

Introduction

While algorithmically-curated media content has much to offer, public concern is mounting about possible negative impacts on individuals and our society as a whole. Most of the focus thus far has been on social media; in comparison, web search has received less attention, though recent studies have found web users are more likely to find and trust news through search than social media sites (Newman 2018). Web search is especially critical in the context of politics, where research has shown it to be among the most commonly-used technologies for finding political information (Dutton et al. 2017).

Prior work has identified that differences in the way search results are presented, in particular their ordering, has substantial effects on user perceptions of content credibility and quality (Pan et al. 2007). These effects may influence users' information-gathering and opinion-formation process substantially enough to impact the outcomes of close elections (Epstein and Robertson 2015). We present an empirical analysis connecting election candidate Google search results and election outcome. Using our data—over 5 million URLs from search results for nearly 4,000 candidates—we find that search media reflect trends theorized in political science around incumbency and election success. In particular, we find that candidates who won general elections in 2018 had less partisan search media than those who lost.

Method & Data

Leading up to the 2018 midterm elections, we scraped the results of Google searches, using as keywords the names of legislative candidates running for office at the federal level (House and Senate). We focus our study on Google as it almost completely dominates the U.S. web search market with a market share of over 90%. We analyze these data by looking at the degree of partisan bias associated with search results' sources, according to partisan attention scores compiled by a team at Harvard's Berkman Klein Center (Faris et al. 2017). This dataset, which covers May 29 through the election on November 6, 2018, collects data using queries for each candidate's name and state abbreviation (i.e., "Dianne Feinstein CA").

Scraping

We used five scrapers to collect this data, each with its own IP address. Each of the five scrapers collects the first page of Google search results for a fifth of the 3,383 candidates daily. We add a depersonalization parameter (“pws=0”) to the end of each query URL in order to avoid any history-based or other personalization. Recent work investigating personalization in political web search has also found that personalization has “little impact” on such queries (Robertson, Lazer, and Wilson 2018).

Partisanship Scores

We use the *partisan attention scores* for 5,798 media sources compiled by the Berkman Klein study as part of their 2017 report to annotate our datasets for the degree of partisan bias they display. These scores, which are expressed on a -1.0 to 1.0 scale, with -1.0 representing extremely left-leaning and 1.0 extremely right-leaning, are generated based on the frequency of media source sharing among over 30,000 users who retweeted either of the two general election candidates (@donaldjtrump and @hillaryclinton), but very rarely retweeted both. We compute the *partisan intensity score* of a page of search results by first taking the absolute value of the partisan attention score of each source on the page, and then averaging all scores on each page of results.

Results

Incumbency Status

Political science literature predicts that incumbents should be more moderate and more centrist in their positions than challengers, in whose best interest it is to be more extreme in order to appeal to the fringe of the party (Groseclose 2001). Extending this theory to the domain of search media, we might expect that incumbents, who are generally more centrist, should receive more mainstream media attention, relative to challengers whose campaigns and subsequent coverage by search media should exhibit higher levels of partisanship. Such theory does not predict what patterns for open seat candidate (those elections without an incumbent) might reflect. We find that incumbents’ web search display lower levels of partisanship than challengers. Interestingly, search media for open seat candidates tracks with challengers, suggesting that candidates running for open seats all find themselves in an uphill battle akin to running against an incumbent.

Partisan Intensity Scores Relative to 2018 Primary Elections

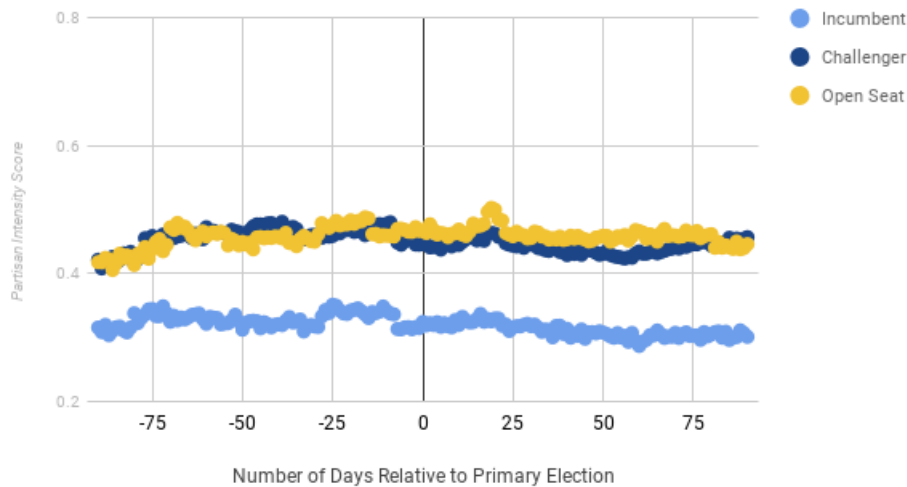


Figure 1. Incumbents' search media are significantly less partisan relative to challengers and those running for open seats.

Election Outcomes

Political science theory that suggests more moderate, centrist positions lead to election victories, due to the importance of garnering a broad voting coalition (Groseclose 2001). We see the same trend in our data: those candidates with lower levels of search media partisan intensity are more competitive. In fact, we are able to use search media partisanship over time to predict election outcomes. We average candidates' partisan intensity scores over the three months prior to the 2018 general elections, and predict that, for each race, the candidate with less partisan search media in that time range will win the election. Using this very simple heuristic, our predictions are 68.4% accurate (compared to less than 50% if we were guessing at chance, since general elections in the U.S. two-party dominant system have at least two major candidates).

Conclusion

In our findings we see evidence that search media provide empirical support for some previously theorized trends regarding incumbency status and election outcome. This work suggests that studying search media and other algorithmically-generated content can provide valuable insight into real-world political phenomena. We hope to extend this work from its current examination of the production of search media to study its consumption: the effect political search results have on users.

References

Dutton, W. H., Reisdorf, B. C., Dubois, E., & Blank, G. (2017). Social Shaping of the Politics of Internet Search and Networking: Moving Beyond Filter Bubbles, Echo Chambers, and Fake News. *SSRN Electronic Journal*. doi:10.2139/ssrn.2944191

Epstein, R., & Robertson, R. E. (2015). The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences*, 112(33). doi:10.1073/pnas.1419828112

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 for Internet & Society Research Paper. <http://nrs.harvard.edu/urn-3:HUL.InstRepos:3375925>

Groseclose, T. (2001). A Model of Candidate Location When One Candidate Has a Valence Advantage. *American Journal of Political Science*, 45(4), 862. doi:10.2307/2669329

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google We Trust: Users' Decisions on Rank, Position, and Relevance. *Journal of Computer-Mediated Communication*, 12(3), 801-823. doi:10.1111/j.1083-6101.2007.00351.x

Robertson, R. E., Lazer, D., & Wilson, C. (2018). Auditing the Personalization and Composition of Politically-Related Search Engine Results Pages. *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW 18*. doi:10.1145/3178876.3186143

FOLLOWER FACTORIES, FALSE AMPLIFIERS AND FAKES FOR SALE: THE POLITICAL ECONOMY OF SOCIAL MEDIA MANIPULATION

Alexander Hogan

Oxford Internet Institute, University of Oxford and Etic Lab

Lisa-Maria Neudert

Oxford Internet Institute, University of Oxford

Introduction

The malicious use of social media platforms tasked with the manipulation of political processes has emerged as critical public interest issue. Globally, political and non-political actors deploy campaigns to sow division, erode trust and orchestrate consensus (Bradshaw, Neudert & Howard, 2018). State-of-the-art information operations rely on an amalgam of automation and big data to disseminate deceptive messages over social media. They use bots, fake personas and sock puppets to game algorithms, inflate opinions and distort discourse – and these instruments of manipulation are for sale. Online market places on both the open and dark web offer deceptive tools, like fake accounts and positive reviews, within the click of a few buttons to anyone with a credit card (Hegelich, 2016).

So far, scholarly work into the initiators of influence campaigns has focused on different actor groups—primarily, Russia’s Internet Research Agency, the alt-right and the military-industrial complex (Benkler, Faris & Roberts, 2018). Profit-driven actors, as opposed to politically motivated ones have received less scholarly attention. It is widely recognized that there are economic incentives to information campaigns, many of which are related to the monetizability of content through advertising. Yet little research has explored how profit-driven actors use online marketplaces to offer products and services that are designed to manipulate as a business model. Using data from the Darknet Market Archives and purposeful keyword search the open and dark web, this paper explores the political economy of digital marketplaces for social media manipulation and examines how profit-driven actors market the tools for the malicious manipulation of social media. Providing an analysis of these tools, their availability, functionality and pricing over time, this paper develops a grounded typology of manipulation marketplaces and hopes to offer insights into their scope and dynamics.

Literature Review

The rise of malicious information operations has appeared as a pressing concern on the global public agenda. In the aftermath of foreign interference in the 2016 US elections, both scholarly and public research have produced a growing body of evidence about influence operations on social media platforms. These inquiries have focused on tracking “computational propaganda” across different platforms and geo-political contexts; social media algorithms as the technical infrastructures that afford such campaigns; and the political actors that execute information operation playbooks (Tucker et al, 2018; Woolley & Howard, 2018). Overwhelmingly, these findings have underscored the role of automated and fake accounts for amplifying viewpoints and manufacturing consensus. Yet, how these tools are developed and acquired by malicious actors has widely escaped scholarly attention.

As social networks have come under attack for offering a fertile ground for influence operations of all colour, academic inquiry has pointed to systemic design flaws that make platforms vulnerable to manipulation. Wu (2017) argues that social platforms capitalize attention through advertising and hence algorithmically promote attention-grabbing content. The more clicks, likes or shares a piece of content has the more visible it will appear. Marwick & Lewis (2016) find that platforms are prone to “attention-hacking”, whereby these mechanisms are gamed with misleading content and false amplification through automated and fake accounts. Similarly, Tucker et al. (2018) argue that social media platforms favor attention regardless of veracity.

By rewarding attention with visibility social networks have created powerful economic and political incentives for algorithm gaming through fake accounts, bots and other forms of content amplifiers. Coupled with a growing demand from malicious actors, social media manipulation has emerged as profitable business strategy that hundreds of digital vendors cater to. Already, phenomenological instances of profit-driven “follower factories” have been reported on (Confessore et al., 2018). As hostile actors continue to look for ways to undermine public discourse, it is pivotal to understand how online marketplaces sustain manipulation and cater to market demands. This paper asks the following questions: (1) What is the scope of digital manipulation marketplaces and their products? (2) In what way do the affordance structures of their products support political manipulation and how do they interact with platform ecosystems? (3) How can platform and policymakers address these issues?

Methodology

This paper analyzes data from two main sources. Firstly, the Gwern Dark Market Archives (Branwen et al., 2015) that comprehensively compile vendor pages and feedback forums from dark net markets from 2013 to 2015. Secondly, we searched the top three search engines on the open web – Google, Yahoo!, Bing – and dark web– Torch, Ahmia, ParaZite– to identify marketplaces with purposefully selected key terms. Modeling our analysis after Joyce, Antonio & Howard (2013) we used crawlers and qualitative content analysis to build a coded spreadsheet of specific variables of interest. These included: available products and services e.g. bots and fake accounts; pricing; currency; social networks targeted; feedback from users about feasibility; targeting options; audience size; and geo-location of accounts used. We will conduct fundamental statistical analysis to show patterns and trends.

To develop a better understanding of these digital marketplaces we set out to perform an iterative typology building process. Typology building is one of the most foundational tasks in political research and is especially important when it comes to investigating new phenomena. (Aronovitch, 2012). Following our sampling strategy our initial dataset identified more than 200 distinct marketplaces. Four main categories were identified: (1) Influence Inflaters that sell simple metric boosters for likes, shares, retweets etc. (2) Fake Factories that offer pre-aged accounts, sock puppets and bots that clients populate with content (3) Human Amplifiers that sell social media engagement from genuine human accounts, often held by users in the global south (4) All-round Agencies that promise custom-tailored influence campaigns using a canon of techniques.

The findings of this analysis will inform a discussion about the role of the emergent political economy for sustaining information operations and provide rich evidence to the current public and scholarly debate on impacts and countermeasures.

References

Aronovitch, H. (2012). Interpreting Weber's Ideal-Types. *Philosophy of the Social Sciences*, 42(3), 356–369. <https://doi.org/10.1177/0048393111408779>

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press.

Bradshaw, S., Neudert, L. M., & Howard, P. N. (2019). *Government Responses To Malicious Uses of Social Media (Working Paper 2019.2)*. Oxford, UK: Oxford University, Project on Computational Propaganda.

Confessore, N., Dance, G. J. X., Harris, R., & Hansen, M. (2018, January 27). The Follower Factory. *The New York Times*. <https://www.nytimes.com/interactive/2018/01/27/technology/social-media-bots.html>

Branwen, G., Christin, N., Décary-Héту, D., Munksgaard, R., StExo, A., Anonymous ... Goode, S. (2015). *Dark Net Market archives, 2011-2015*. <https://www.gwern.net/DNM-archives>

Hegelich, S. (2016). *Social Bots: Invasion der Meinungsroboter*. Konrad-Adenauer-Stiftung, 221, 1–9. <http://www.kas.de/wf/de/33.46486/>

Joyce, M., Antonio, R., & Howard, P. N. (2013). *Global Digital Activism Data Set*. ICPSR. <http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/34625/version/2>

Marwick, A., & Lewis, R. (2017). *Media Manipulation and Disinformation Online*. Data & Society Research Institute. https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

Tucker, J. A., Guess, A., Barbera, P., Vaccari, C., Siegel, A., Sanovich, S., ... Nyhan, B. (2018). *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature*. *SSRN Electronic Journal*, 95. <https://doi.org/10.2139/ssrn.3144139>
Woolley, S., & Howard, P. (2018). *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*. Oxford: Oxford University Press.

Wu, T. (2017). *The Attention Merchants: The Epic Struggle to Get Inside Our Heads*. Atlantic Books.