

UCLA

UCLA Previously Published Works

Title

Unmasking Hate on Twitter: Disrupting Anonymity by Tracking Trolls

Permalink

<https://escholarship.org/uc/item/4mm6850p>

Author

Ascher, Diana L

Publication Date

2019-04-25

Peer reviewed

Abstract

Notions of free speech and expectations of speaker anonymity are instrumental aspects of online information practice in the United States, which manifest in greater protections for speakers of hate, while making targets of trolling and hate speech more vulnerable. In this chapter, we argue that corporate digital media platforms moderate and manage “free speech” in ways that disproportionately harm vulnerable populations. After being targets of racist and misogynist trolling ourselves, we investigated whether new modes of analysis could identify and strengthen the ties between the online personas of anonymous speakers of hate and their identities in real life, which may present opportunities for intervention to arrest online hate speech, or at least make speakers known to those who are targets or recipients of their speech.

Keywords: digital media, hate speech, information studies, racism, social media networks, Twitter

10

Unmasking Hate on Twitter

Disrupting Anonymity by Tracking Trolls

Diana L. Ascher

1. Introduction

Of primary interest in this chapter is the apparent emboldening of neo-nazi¹ hate speech, the implications of this phenomenon for vulnerable populations, and potential modes of remedy. Our recent experience with neo-nazi hate group members who actively engage in social media trolling led us to think about the implications of protected hate speech, and the ways in which digital media platforms protect the anonymity of speakers, while making it nearly impossible for the targets of hate speech to know its origin. Whereas Ku Klux Klan (KKK) members in the analog era used robes and hoods to assume a state of pseudonymity, certain information practices have created an emboldened sense of righteousness among neo-nazis, a desensitization of the general public to hate speech, and an exacerbation of the precariousness of the most vulnerable members of society.

Many of the protections afforded to speakers of hate on the internet are governed by legal decisions. Courts are increasingly forced to rule on “true threat” cases to determine the degree to which online comments constitute a threat that can lead to violence or other types of harm (Best 2016). Federal legislation limiting speech and information practice on social networking platforms remains unclear, though states are gaining some traction in criminalizing the harms caused by revenge porn and other malicious online communications. Constitutionally protected speech is a major sticking point when attempting to adjudicate the kinds of speech that occur online (Williams 2014), often leaving victims of misogynist, racist, homophobic, and other forms of persecution speech with little legal recourse or protection. The Communications Decency Act (CDA), which grants protections and immunity from prosecution to technology companies for content posted to their online platforms, presents an even greater challenge for victims of anonymous trolls and/or hate speech in social media networks. The Act characterizes technology companies as “vessels” for content, with no accountability for the propagation of messages through their networks. This lack of

accountability is counterintuitive to those of us who know that the algorithmic curation and circulation of content through social media are tied directly to algorithmic advertising mechanisms and decision-making by human commercial content moderators (Roberts 2016; Noble 2018).

While platforms—such as Facebook, Twitter, Tumblr, and Instagram—may eschew any responsibility for hate speech content, or grapple to figure out the limits of speech that might invoke harm (Huff 2017), we contend that other avenues of protection from anonymous trolling might empower victims of targeted hate speech in social media networks. We present an account of how we tracked down the true identities of members of a neo-nazi hate group on Twitter to stimulate a conversation about the tension between free speech and criminalizing hate speech, and to determine whether a de-anonymizing toolkit for victims of hate speech on Twitter is a worthy endeavor.

2. Free Speech, Power, and Anonymity Online

Online, anonymity means that an author's identity is unknown. Sometimes, this comes in the form of pseudonymity, in which a message is attributed to an online persona, represented by a name,

also called a handle, other than the author's. Pseudonymity can be insulating for authors who are more confident sharing their messages when their identities are unknown, and it is in this context that the internet has offered a unique space for people who share an interest to connect in a semi-protected environment. In other contexts, pseudonymity may represent an author's desire to compartmentalize the types of messages shared. For example, works by Mark Twain are differentiated from those attributed to Samuel Clemens, though both were penned by the same man. Thus, pseudonymity affords an author a measure of identity, but in the form of an alternate persona. This can be beneficial for social media contributors who wish to build a following that is not directly connected with their personal lives. One celebrity who has perfected the art of such compartmentalization is Beyoncé. The performer's Instagram account has 109 million followers, who may perceive an authentic connection to the artist; however, Beyoncé's public identity is the creation of a carefully crafted brand strategy that reveals little of the particulars of daily life. You never see indicators of time or place in the images that populate her social media posts; all images, like her music, are served up

with a consistency of message and within the boundaries she has established since wresting control of her publicity from her manager father in 2011. Beyoncé controls what her fans know about her personal life, with very few exceptions. And recently, she has leveraged social media to make explicit her ideological and political stances, despite a conscious decision to decline personal interviews since mid-2013. Her revelations are conveyed through her art. Consider this accomplishment in contrast to internet celebrities, such as the Kardashian family, whose self-promotion of sex tapes and coverage of controversial behavior in increasingly personal contexts have launched a media empire that extends from social media to mass media.

Such personas can be deployed for both good and ill, of course. Certainly, the recent revelations about Macedonian fake news efforts represent the deceptive potential of pseudonymity (Subramanian 2017). Purveyors of hate speech on Twitter exploit a false sense of security that users have in their anonymity while accumulating social power under the guise of pseudonyms. What is it about Twitter that makes its users so vulnerable to hate speech? Part of the appeal of social media platforms for members of

marginalized groups is that the networks that form among users have the potential to connect individuals with others who share some interest, despite the constraints of space and time. For example, the internet has enabled online support networks among people interested in rare diseases; such connections had not been possible prior to the ability to search for others worldwide. The cultural phenomenon known as “Black Twitter” is another such case, this time of African Americans using the platform to communicate, signify, and organize responses and resistance to racialized oppression (Brock 2012). Of course, when people are looking for collaborators and/or commiserators online, they also make themselves vulnerable to users of the platform who, for a host of reasons beyond the scope of this chapter, are willing to invest in trolling them.

3. The Internet *Does* Know You’re a Dog

It is important to understand how anonymity works in social media networks if we wish to think differently about combatting hate speech on these platforms. Depending on the platform, members of social media networks may engage with one another with varying degrees of anonymity. One of the most demonstrative examples of

the misperception of anonymity the general public associates with the internet, a cartoon by Peter Steiner published by *The New Yorker* in 1993, has come to represent the information practices that shape online identity (see Figure 10.1). The cartoon is among the first and most enduring memes to characterize the online world for the general public, when the promise of the internet as a democratizing technology was the prevailing perspective of the time. Sherry Turkle (1995, 184-5), among others (e.g., Rheingold 1993; Negroponte 1995), heralded the internet's inherent anonymity as a democratizing force. Turkle famously offered one gamer's characterization of what it's like to engage online:

You can be whoever you want to be. You can completely redefine yourself if you want. You don't have to worry about the slots other people put you in as much. They don't look at your body and make assumptions. They don't hear your accent and make assumptions. All they see are your words.

And while privacy concerns were raised with respect to online identity, more attention was directed to problematizing online addiction than to the potential for disproportionate anonymity and

protections (e.g., Negroponte 1995). The misperception of online anonymity was firmly entrenched in American media culture, as shown in [Figure 10.1](#):

[INSERT FIGURE 10.1 HERE]

By the early 2000s, digital media scholars had debunked these ideas that the body could be liberated from the online experience and showed how patterns of online interaction are always racialized and gendered, much in the same ways they are offline (Nakamura, 2002). Jessie Daniels (2009) wrote one of the most important monographs describing how white supremacy and racist organizations work online, showing how white supremacist groups use the web to bolster themselves through both cloaked websites that mask their hate speech in seemingly credible or legitimate mainstream websites and overt racist speech and websites used for that sole purpose. More recently, while some strides have been made in information literacy with respect to fraudulent online identities—perhaps most effectively as a result of the MTV program, *Catfish*—public awareness of the power dynamics inherent in online anonymity remains low.

Lisa Nakamura describes the performative nature of online identities, in which a user plays the role of an individual of a particular gender and race, engaging in what she calls identity tourism. Describing the scene in Figure 10.1, Nakamura explains that the dog avails itself of “the freedom to ‘pass’ as part of a privileged group, i.e. human computer users with access to the Internet. This is possible because of the discursive dynamic of the Internet,” particularly on platforms that permit creation of a user identity without a verifiable email address (Nakamura 2002, 1). In online gamespaces, users frequently employ identities as characters. However, in the domain of social media platforms, identities are not considered characters or roles in the same manner. Identity on social media platforms is a self-representation, tailored for the specific network audience. Moreover, one’s Twitter identity can be seen to take on an ideological dimension, as endorsements and redistribution of preferred content—to the exclusion of less-preferred content—signify facets of the user’s belief system (Ascher 2014, 2017; Brock 2012).

Thus, social media platforms create specific expectations of anonymity through their user engagement policies. While network

members believe they are somehow protected from persecution by virtue of platform-dependent anonymity, the technological expertise residing in hate groups creates a significant danger for members of vulnerable populations. Trolls understand that no one is anonymous online. And while the tech-savvy white nationalists have the wherewithal to de-anonymize members of vulnerable groups and target them with hate speech, the layperson is ill-equipped to employ technological protections or to use technology to unmask assailants in any useful way. Furthermore, the recent emboldening of white nationalists and other hate speakers online seems to negate the benefits of unmasking, in terms of social proof.

We note that targets in systems of white supremacy and racial categorization are marked, both online and otherwise in real life, because participants in open commercial media platforms are never truly anonymous. The moment a single marker is triggered that indicates a user is not part of the dominant cultural norm in a platform, the differentiating trait becomes a trolling target, and these traits are often expressed as racialized, gendered, and sexual orientation markers. In the analog era, there was no shortage of

neo-nazi rhetoric and propaganda. Messages of hate came in a variety of forms, making use of every communication medium available. However, the investment of time and labor to create and disseminate neo-nazi hate speech was considerably greater in the pre-internet era than it is today. Furthermore, the material connection between hate speech and its effects on vulnerable populations was easier to trace.

4. Neo-nazi Hate Speech Online

Our speculation about how social media platforms provide protections to trolls and misconceptions about anonymity to vulnerable communities is grounded in our experience with a white nationalist group based in Southern California, informed by Social Proof Theory, and approached using network visualization and social network analysis.

Social Proof Theory (sometimes called informational social influence) is one of six principles of persuasion advanced by Robert Cialdini (1993), which describes the tendency of people to perform certain actions when they identify with other people who performed those actions previously (Cialdini 1993). In other words, the theory posits that individuals are biased toward

following the crowd—people assume that “if many similar others are acting or have been acting in a particular way within a situation, it is likely to represent a good choice” (Cialdini 2009). The underlying logic of the theory rests in the assumption that others have special knowledge about the situation, which the individual decision maker lacks. Thus, individuals fail to question the propriety of the behaviors modeled by others in context. By comparing their behavior with referent others, individuals validate the “correctness” of their opinions and decisions (Festinger, 1954). Research has demonstrated the role of social proof in a wide variety of activities, including approaching a frightening dog, deciding whether and how to commit suicide, donating funds to charity, engaging in promiscuous sexual activity in a “safe” versus “unsafe” manner, littering in a public place, and returning a lost wallet (Cialdini et al. 1999).

In sum, we can think of Social Proof Theory as a motivation for the popular adage: *Birds of a feather flock together*. This phenomenon is particularly relevant to information cascades in social media networks, in which individuals undertake specific information practices to signal ideological alignment. Social proof

may be observed in hashtag use on Twitter, which communicates the user's perspective on a topic. For example, Twitter users convey solidarity with individuals who have been targets of sexual harassment and abuse by tweeting #MeToo. Social proof also may be observed in social media networks in the language and tone employed by individuals who identify with the white nationalist movement. We assert that specific information practices of President Trump have influenced how those who identify with this movement conceptualize appropriate information practice in the context of social media networks.

President Trump's redistribution of content from the Twitter accounts of neo-nazi leaders serves as a legitimization signal and facilitates an emboldening of white supremacists online. De-anonymizing typically happens at the direction and in the service of those in power, such as the NSA, FBI, law enforcement, or university administrators. It rarely occurs to the benefit of the victims of hate speech online. Often, this disparity stems from corporate platforms that, as a matter of policy, do not de-anonymize white supremacists and trolls who propagate hate under the guise of anonymity.

We can look to similar mechanisms of anonymity offline for insight into the advantages and disadvantages of anonymous or pseudonymous communication of hate online. For more than a century, members of the KKK have hidden their identities under hoods and robes that symbolically convey hatred and threaten harm. However, the individuals donning KKK regalia often are known by their victims and by the community at large. As a teen growing up in the South in the late 1980s, I (Ascher) witnessed the overt intimidation and symbolic communication of threat when the KKK marched through my high school. I recall vividly the story of Louis Kittler, a Jewish cobbler in a small North Carolina town, who could identify Klansmen by their shoes. Even under the ostensible guise of hood and robe, identities were no secret. The KKK garb provided a means for the community to cling to plausible deniability, as members looked the other way, permitting the symbolic intimidation. What, then, was the function of the racist garb?

We can conceive of myriad ways in which the KKK hood protects the wearer and harms the target. The hood is a means for other members of the community to deny complicity, just as online

pseudonymity and invoking freedom of speech makes it easy for members of the community to avoid getting involved. The KKK robe and hood are material forms of social proof, used to reify the racist patriarchal social order. Permissive pseudonymity bolsters misperceptions about the security of users' personal information and provides a substitute for social proof, which, otherwise, might hold individuals accountable for their online behavior. This is precisely how technology companies shirk responsibility for enforcing standards of conduct on their social media platforms.

The online pseudonym was once a guiding light of internet culture, a crucial protection for whistleblowers and communities with a legitimate fear of being exposed. Now, it's increasingly seen as a threat. Worse, it seems more and more likely that platforms will respond to Russia concerns by tightening restrictions on online anonymity, and driving webgoers to live more and more of their online life under legal names (Brandom, 2017).

An interesting side effect of the Twitter platform is its contributing to the emboldening of trolls online and in the real world. Thanks to

people such as Kim Kardashian and the decline of scripted television, social media has become a means for some to attain celebrity status. Two aspects are notable. First, since inflammatory content draws more attention than uncontroversial topics, the general public has become desensitized to derogatory language. This is not surprising; however, algorithmic sensationalism amplifies derogatory messages in social media networks (Ascher 2017). Second, the line between Hollywood and reality has blurred. In the year since the 2016 U.S. presidential election, we have seen not only the emboldening of trolls on Twitter, but also their rise to social media celebrity status—using their true identities. For example, Milo Yiannopoulos, who earned the distinction of receiving a lifetime ban from Twitter for his role in inciting harassment against *Saturday Night Live* actress Leslie Jones, has inspired violent protests on college campuses where he has been invited to speak (Rakhim 2017). Thus, the lever of social pressure that usually serves to discourage blatant hate crimes offline has transformed into a sort of twisted notoriety—the sort that demands five-figure speaking fees.

5. Exploring De-anonymizing Tools

When a white nationalist group blanketed the UCLA campus with racist and anti-Semitic flyers in the spring of 2017 (see [Figure 10.2](#)) and targeted one of our faculty with online hate speech and threats, we conducted a social network analysis to learn about the group and assess the risk of harm its members represented. Our purpose was twofold: assess the immediate risk, and determine whether a toolkit for de-anonymizing purveyors of online hate is a worthy endeavor.

[INSERT FIGURE 10.2 HERE]

Our investigation was not simple. As information studies researchers, we are experienced with a variety of techniques to uncover the origins of and modifications to electronic documents, including text, photos, and videos. Many of the techniques involve tracing metadata—data that describe the information—and geospatial data to identify the creator(s) and/or individual(s) responsible for modifying digital information. Stripping metadata from electronic documents is not difficult, but it does require conscientious effort on the part of the person who posts the content online. Usually, a skilled researcher can track down the source of, say, a photo of a group of neo-nazis posing together at a graffiti-

covered crematorium, with little difficulty (see [Figure 10.3](#)).

However, we found the level of technical sophistication demonstrated by the group members, who call themselves the Beach Goys, challenging.

[INSERT FIGURE 10.3 HERE]

Tracing the email address on the flyers and the handles of the trolls who were threatening the faculty member, we found the group's pseudonymous Twitter handle: @BeachGoys. Using social network visualization application NodeXL Pro, we imported a list of all Twitter users who engaged with @BeachGoys, and graphed the relationships in the network to learn more about their activities (see [Figure 10.4](#)).

[INSERT FIGURE 10.4 HERE]

Armed with a list of individual Twitter handles and a sense of the connections among the users, we eventually located several members of the Beach Goys based in Long Beach and the greater Los Angeles area. Members of this group periodically get together for hikes to West Coast sites that are said to be the locations selected by Hitler for nazi occupation and continuation of the extermination of the Jewish people. Photos taken on these hikes

depict several men in their 20s and 30s, decked out in hiking gear (see [Figure 10.5](#)). Faces in the images posted by the group are masked by the superimposition of several cartoon character heads, including Pepe, a frog appropriated from children's book illustrator Matt Furie, which has become a symbol of racist hate (Hunt, 2017). Metadata had been stripped from these images, suggesting the group members are not only aware of the geolocation and reidentifying power of metadata, but also of the methods for removing it.

Interestingly, by searching for user names associated with the photos—the pseudonyms by which group members are known on Twitter and other platforms—we learned that the self-proclaimed leader of the group participates in a Frisbee golf league, which enabled the identification of several group members.

[INSERT FIGURE 10.5 HERE]

Using Google image search, we located several similar photos posted to various online forums, including documentation on the Daily Stormer, a neo-nazi website, of the group attending a talk by retired California State University, Long Beach Professor Kevin MacDonald (see [Figure 10.6](#)), who is known for an anti-

Semitic trilogy that argues anti-Semitism is a rational reaction to Jews' genetic predisposition for outcompeting the white Christian creators of Western civilization (Southern Poverty Law Center, 2018; Taylor, 2016).

[INSERT FIGURE 10.6 HERE]

In the same comment thread (Los Angeles Beach Goys—Book Club—Daily Stormer BBS), we found a conversation among Beach Goys members and individuals inquiring about joining the group. In this conversation, prospective members ask about the vetting protocol, which begins with a Skype verification session. A member using the handle Salad Snake describes the group as being based in Los Angeles, with members living as far away as the San Gabriel Valley and the Inland Empire. He alludes to coordination with other white nationalist groups, mentioning meetups in Irvine that are co-sponsored with a group from San Diego. Salad Snake explains:

The way we do it is we have you, the other founding members, and I talk on a skype call. This gives us all a chance to see if it's a good fit. We're normal guys. We jave [sic] jobs, bills, girlfriends

etc. We'll joke about jews for a bit and talk about the group. The thing that brings us all together is we are Nationalists for people of European heritage.

Everything else we believe supports this main premise ie: traditionalism; the institutions that make our civilization strong and healthy etc. We are pro white, Pro West, and having a great time about it.

How do you feel about that?

Since Trump's election, a few of these group members have spoken openly with reporters about their vision for "purification" (e.g., Bhattacharya, 2016; New Yorker, 2016). Thus, these group members who were careful to strip the identifying metadata and conceal their faces with superimposed Pepe the Frog images, suddenly became willing to be named in the popular press. What accounts for this change—effectively, moving from the shadows into the limelight—is a sense that the power dynamic has shifted.

Several overt actions communicated this sea change in what may and may not be voiced openly. First, partisan news media, coupled with the "fake news" revelations of 2017, created filter bubbles (Pariser 2011, 2012) unlike any experienced

previously. As a point of reference, the Brookings Institution notes that the “20 largest fake news stories of the 2016 election generated 1.3 million more social media engagements than the top 20 real news stories” and Americans’ trust in the mass media “to report the news fully, accurately and fairly” last year dropped to an all-time low of 32% (see [Figure 10.7](#)) (Swift 2016; West 2017).

[INSERT FIGURE 10.7 HERE]

Compounding this phenomenon, President Trump tweeted reactions and affirmations of news stories in lockstep with certain news broadcasts (Kludt and Yellin, 2017), and declared other mainstream news media operations to be purveyors of so-called fake news. In addition, Trump blocked access to his tweets for journalists and other Twitter users who expressed disagreement with his statements. An ongoing federal lawsuit charges that Trump’s practice of blocking critics from his personal Twitter account is a violation of the First Amendment (Schonfeld, 2017). Simultaneously, a new form of yellow journalism— *algorithmic sensationalism*—has arisen from information practices at news organizations that disproportionately amplify inflammatory content and lack a mechanism for applying timely human judgment

(Ascher, 2017). All of these actions have undermined public trust in media content authenticity and veracity.

Trump reinforces and publicizes his connections with neo-nazis through copying and pasting the content of tweets originating from Twitter accounts using known neo-nazi pseudonyms. This information practice, coupled with enthusiastic congratulations from infamous former KKK leader David Duke for sharing videos that show what appear to be Muslim men destroying Christian relics and assaulting non-Muslim men (Giaritelli, 2017), has emboldened neo-nazi social media users (see [Figure 10.8](#)).

[INSERT FIGURE 10.8 HERE]

Furthermore, Trump includes language in his tweet diatribes and in his formal speeches that echo neo-nazi propaganda. For example, as concerns mounted about the potential for violence at rallies, Trump reinforced the delicate religious reframing that neo-nazis employ to protect their right to assemble and speak. His tweets often are accompanied by memes drawn from World War II-era propaganda imagery, as shown in [Figure 10.9](#). These actions convey an alignment with the chosen reframing, and are symbolic of tacit support for the neo-nazi agenda. Furthermore, Trump's

initial refusal to condemn and his subsequent watered-down denunciation of neo-nazi hate crimes in Charlottesville and Boston provide just the right amount of wiggle room necessary to prevaricate actions that would have been condemned a mere 12 months earlier.

[INSERT FIGURE 10.9 HERE]

6. Conclusion: Unmasking Online Speakers of Hate

That there seemingly is no shame associated with voicing racist, sexist, homophobic, and misogynist opinions in public discourse should be of concern to everyone. While the pseudonymity of Twitter helps users with similar opinions find one another and reinforces their notions of community, inundation of inflammatory opinions on social media platforms contributes to a dangerous social desensitization to harmful rhetoric and blatant fake news. What this means for anonymity in social media networks is that those in power use pseudonymity to their advantage and are unconcerned with being exposed, while vulnerable members of persecuted groups increasingly depend heavily on the assumed

protections of anonymity, even as these protections are challenged by the ruling administration. This transposition of the need for anonymity in social networks shifts the chilling effect from neo-nazis and other hate groups who previously went to great lengths to protect their identities, to the vulnerable and historically unprotected, marginalized members of society.

We were fortunate to have determined the true identities of several members of the Beach Goys, which provided some potential recourse had the trolls escalated their threats against UCLA faculty. However, we are alarmed by the notoriety and emboldening of these groups, particularly as they have been legitimized and propagated by President Trump's information practice. While we are troubled by the emboldening of hate groups and the subsequent chilling effect experienced by marginalized communities, we note that those in power typically make it easier for citizens in ideological agreement with them to speak openly, support one another, and act in their own interests. Of course, this necessarily makes it harder for those who oppose the ideology of those in power to communicate openly, work together, and effect change.

In August and September of 2017, members of the Beach Goys attended meetings of the Santa Monica Committee for Racial Justice (see [Figure 10.10](#)).

[INSERT FIGURE 10.10 HERE]

This development underscores the imperative that opposition groups working on behalf of human and civil rights must retain the right and ability to exchange information anonymously, yet have means to counter the effects of social proof observed in the information practice of purveyors of hate speech online through de-anonymization. We note, in conclusion, that the Federal Bureau of Investigation (FBI) has expressed grave concern over the quiet radicalization of white men, and reported that the single largest terrorist threat is domestic white supremacists infiltrating law enforcement (Speri, 2017). Additionally, the Department of Homeland Security (2009) reports concern “that rightwing extremists will attempt to recruit and radicalize returning veterans in order to boost their violent capabilities.” Anonymity online, particularly among speakers of hate and trolls, makes it difficult for municipalities and the public to hold such domestic terrorists accountable for their intimidation and threats.

We believe, based on our study of the cloaking protections invoked by the white supremacists we studied, that social media platforms must work with researchers and policymakers to protect free speech online, yet provide recourse for victims and targets of hate speech. Topics for future discussion and research include new forms of social proof in social media networks, social media literacy with respect to anonymity and privacy online, and alternative means of exposing the true identities of network members who threaten and torment others with racist, sexist, homophobic, misogynist hate speech.

Notes

References

Ascher, D. L. 2014. "A Time Analytic Framework for Information Practice." *2014 ASIS&T SIG -USE Symposium Context in Information Behavior Research*.

Ascher, D. 2017. "The New Yellow Journalism: Examining the Algorithmic Turn in News Organizations' Social Media Activity through the Lens of Cultural Time Orientation." University of California, Los Angeles.

Best, A. J. 2016. “*Elonis v. United States*: The Need to Uphold Individual Rights to Free Speech while Protecting Victims of Online True Threats.” *Maryland Law Review* 75 (4): 1127–58.

Bhattacharya, S. 2016. “‘Call Me a Racist, but Don’t Say I’m a Buddhist’: America’s Alt Right.” *The Observer*, October 9.
<https://www.theguardian.com/world/2016/oct/09/call-me-a-racist-but-dont-say-im-a-buddhist-meet-the-alt-right>.

Brandom, R. 2017. “Russia’s Interference Could Be the End of Social Media Anonymity.” *The Verge*, November 1.
<https://www.theverge.com/2017/11/1/16592374/russia-facebook-ads-clint-watts-social-media-anonymity-privacy>.

Brock, A. 2012. “From the Blackhand Side: Twitter as a Cultural Conversation.” *Journal of Broadcasting & Electronic Media* 56 (4): 529–49.
doi:10.1080/08838151.2012.732147.

Brown, A. 2017. “Rogue Twitter Accounts Fight to Preserve the Voice of Government Science.” *The Intercept*, March 11.
<https://theintercept.com/2017/03/11/rogue-twitter->

accounts-fight-to-preserve-the-voice-of-government-science.

Cialdini, R. B. 1993. *Influence: The Psychology of Persuasion*. New York: HarperCollins.

Cialdini, R. B., W. Wosinska, D. W. Barrett, J. Butner, and M. Gornik-Durose. 1999. "Compliance with a Request in Two Cultures: The Differential Influence of Social Proof and Commitment/Consistency on Collectivists and Individualists." *Personality and Social Psychology Bulletin* 25 (10): 1242–53. doi:10.1177/0146167299258006.

Cialdini, R. B. 2009. *Influence: Science and Practice*. Boston, MA: Pearson Education.

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

Domonoske, C. 2017. "Twitter Sues Homeland Security to Protect Anonymity of 'Alt Immigration' Account." *NPR*, April 6.

<http://www.npr.org/sections/thetwo->

[way/2017/04/06/522914335/twitter-sues-homeland-](http://www.npr.org/sections/thetwo-way/2017/04/06/522914335/twitter-sues-homeland-)

[security-to-protect-anonymity-of-alt-immigration-account](http://www.npr.org/sections/thetwo-way/2017/04/06/522914335/twitter-sues-homeland-security-to-protect-anonymity-of-alt-immigration-account).

Festinger, L. 1954. "A Theory of Social Comparison Processes."

Human Relations, May.

Fleishman, G. 2000. "Cartoon Captures Spirit of the Internet." *The*

New York Times, December 13.

<http://www.nytimes.com/2000/12/14/technology/cartoon-captures-spirit-of-the-internet.html>.

Giaritelli, A. 2017. "David Duke after Trump Tweets Videos of

Muslims Attacking People: 'That's Why We Love Him!'"

Washington Examiner, November 29.

<http://www.washingtonexaminer.com/david-duke-after-trump-tweets-videos-of-muslims-attacking-people-thats-why-we-love-him/article/2641982>.

Grannan, K. 2017. "The First-Time Voters." *The New Yorker*,

September 7.

<https://www.newyorker.com/magazine/2016/10/31/first-time-voters>.

Huff, L. 2017. "Mark Zuckerberg Vows to Remove Hate Speech

from Facebook." *The Hollywood Reporter*, August 16.

<http://www.hollywoodreporter.com/news/mark-zuckerberg-vows-remove-hate-speech-facebook-1030327>.

Hunt, E. 2017. "Pepe the Frog Creator Kills Off Internet Meme Co-Opted by White Supremacists." *The Guardian*, May 7.

<https://www.theguardian.com/world/2017/may/08/pepe-the-frog-creator-kills-off-internet-meme-co-opted-by-white-supremacists>.

Kittner, S. 2018. "A Trial Is All We Ask . . ." *Sam Kittner Photographer*, Accessed January 19.

http://kittner.com/a_trial_is_all_we_ask.php.

Kludt, T. and T. Yellin. 2017. "Trump Tweets and the TV News Stories behind Them." *CNNMoney*, October 10.

<http://money.cnn.com/interactive/media/trump-tv-tweets/index.html>.

Nakamura, L. 2002. *Cybertypes: Race, Ethnicity, and Identity on the Internet*. New York: Routledge.

Negroponte, N. 1995. *Being Digital*. New York: Alfred A. Knopf.

Negroponte, N. 1996. *Being Digital*. New York: Vintage Books.

Noble, S. U. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

Pariser, E. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin Press.

Pariser, E. 2012. *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. London: Penguin Books.

Pfiffner, J. P. 2017. "The Unusual Presidency of Donald Trump." *Political Insight* 8 (2): 9–11.
doi:10.1177/2041905817726890.

Rahim, Z. 2017. "President Trump Threatens to Yank U.C. Berkeley's Federal Funding over Protests against Milo Yiannopoulos." *Time*, February 2.
<http://time.com/4657688/trump-milo-yiannopoulos-breitbart/>.

Rheingold, H. 1993. *The Virtual Community: Homesteading on the Electronic Frontier*. Cambridge, MA: MIT Press.

Roberts, S.T. (2016). Commercial content moderation: Digital laborers' dirty work. In Noble, S.U. and Tynes, B. (Eds.), *The intersectional internet: Race, sex, class and culture online* (pp. 147-159). New York: Peter Lang.

Schonfeld, Z. 2017. "Donald Trump Blocked Me on Twitter. Should I Sue Him?" *Newsweek*, June 22.

<http://www.newsweek.com/should-i-sue-donald-trump-blocking-me-twitter-knight-foundation-627963>.

Southern Poverty Law Center. 2018. "Kevin MacDonald." *Southern Poverty Law Center*. Accessed January 20.

<https://www.splcenter.org/fighting-hate/extremist-files/individual/kevin-macdonald>.

Speri, A. 2017. "The FBI Has Quietly Investigated White Supremacist Infiltration of Law Enforcement." *The Intercept*, January 31.

<https://theintercept.com/2017/01/31/the-fbi-has-quietly-investigated-white-supremacist-infiltration-of-law-enforcement/>.

Subramanian, S. 2017. “The Macedonian Teens Who Mastered Fake News.” *Wired*, May 1.

<https://www.wired.com/2017/02/veles-macedonia-fake-news/>.

Swift, A. 2016. “Americans’ Trust in Mass Media Sinks to New Low.” *Gallup*, September 14.

<http://news.gallup.com/poll/195542/americans-trust-mass-media-sinks-new-low.aspx>.

Taddeo, M. and L. Floridi. 2015. “The Moral Responsibilities of Online Service Providers.” *Law, Governance and Technology Series The Responsibilities of Online Service Providers* 22 (6): 13–42. doi:10.1007/978-3-319-47852-4_2.

Taylor, J. 2016. “Energized by Trump’s Win, White Nationalists Gather to ‘Change the World.’” *NPR*, November 20.

<https://www.npr.org/2016/11/20/502719871/energized-by-trumps-win-white-nationalists-gather-to-change-the-world>.

Turkle, S. 1995. *Life on the Screen: Identity in the Age of the Internet*. New York: Simon & Schuster Paperbacks.

U.S. Department of Homeland Security. 2009. “Rightwing Extremism: Current Economic and Political Climate Fueling Resurgence in Radicalization and Recruitment. Prepared by the Extremism and Radicalization Branch, Homeland Environment Threat Analysis Division.” *Extremism and Radicalization Branch, Homeland Environment Threat Analysis Division; Declassified, April.* <https://fas.org/irp/eprint/rightwing.pdf>.

West, D. M. 2017. “How to Combat Fake News and Disinformation.” *Brookings*, December 18. https://www.brookings.edu/research/how-to-combat-fake-news-and-disinformation/?utm_medium=social&utm_source=twitter&utm_campaign=gs.

Williams, L. 2014. “California’s Anti-revenge Porn Legislation: Good Intentions, Unconstitutional Result.” *California Legal History* 9: 297–338.

¹ We have deliberately chosen to write neo-nazi without capitals.