Chapter 15. Studying Unknown Unknowns: Lessons from Critical Making on Twitter

Whitney Lew James
University of Notre Dame

In June 2020, SARS-CoV-2, the novel coronavirus that caused a global pandemic, had been known to the medical and scientific community for approximately six months. There were still many unanswered questions about COVID-19, but the American public was highly divided on what little concrete information was available as well as the credibility of information sources. The Pew Research Center reported that between April and June 2020, Republicans and Democrats became increasingly divided on several issues related to the pandemic ("Republicans, Democrats Move Even Further Apart"). Another Pew poll indicated that both Republicans and Democrats felt the CDC and other public health organizations were most likely to get COVID facts correct; however, 54% of Republicans felt Trump and his administration reported accurate information compared to only 9% of Democrats (Mitchell et al.). Why were Americans becoming increasingly divided about what is scientifically factual and who provides accurate information?

While many factors are responsible for divisions among the American public, one likely contributing factor is information echo chambers on online news sources, particularly on social media platforms, where 52% of Americans get most of their news (Suciu). Echo chambers, also known as filter bubbles, occur when individuals are exposed to information and ideas that reinforce their existing views—creating an echo and amplification of their own ideas—while suppressing alternative perspectives (Sunstein). While the existence of echo chambers remains debated in scholarly circles, evidence suggests that both Twitter and Facebook are "dominated by echo chambers" (Cinelli et al. 6). Although studies suggest that the tendency to seek out information that confirms pre-existing opinions is particularly strong in content consumption on online social media (Del Vicario et al.; Garimella et al.), this phenomenon does not only occur in online spaces. Social media echo chambers are simply the latest iteration of the homophily principle, or the human tendency to interact with like-minded peers. Content curation and recommendation algorithms on social media platforms are specifically designed to support homophily, further exacerbating the likelihood of social media echo chambers.

The ways that recommendation algorithms—nonhuman, rhetorical agents that provide personalized recommendations based on aggregated user behavior

DOI: https://doi.org/10.37514/PRA-B.2022.1664.2.15

data—contribute to echo chambers is of great concern to digital rhetoricians. As John Gallagher, Estee Beck, and Annette Vee, among others, have argued, computer code and predictive algorithms are rhetorical agents that must be studied within our field and incorporated into our pedagogy. What makes this type of research incredibly difficult is that both echo chambers and the recommendation algorithms that appear to contribute to their existence are opaque. The predictive algorithms used by recommendation systems are proprietary "products" that social media companies fiercely guard. In this sense, predictive algorithms are "unknown unknowns"—objects of study that researchers do not know what they don't know about (Brunton and Nissenbaum). Studying unknown unknowns is highly difficult and relevant given the increasing influence of predictive algorithms on all facets of life and requires innovative and non-traditional research methods.

In 2019, I conducted a live study of Twitter's recommendation system, Who to Follow or WTF, using a series of Twitter bots, small bits of code that automatically perform specific functions, to understand how recommendation algorithms may contribute to the echo chambers phenomenon. Because I wanted to learn about and critique Twitter, using the platform itself seemed like the best method for accomplishing both goals. I chose to learn about algorithms by creating my own simple algorithm and to display my critique of predictive technology on the very platform I studied. From March to November 2019, I created eight Twitter accounts that used WTF's account recommendations to follow content aligned with a specific lifestyle, ideological, or demographic group. Each of these accounts was then connected with a bot that automatically retweeted the account's feed making the contents of the individual Twitter feeds visible to the public. In essence, I used bot automation to critique recommendation automation. Twitter terminated the project by suspending all the accounts associated with my study—as well as my personal Twitter account and the account of a former employer—for "platform manipulation and spamming." Ironically, the suspension of my access to Twitter gives insight into how the company silences critique while their own algorithms appear to be designed to propagate the uneven spread of information that I was accused of committing. Over the course of the study, I not only learned about recommendation algorithms and echo chambers, but also about the challenges of conducting research on a live social media platform.

This chapter will discuss why and how I used Twitter bots as an activist research method to study predictive algorithms as well as the major obstacles I faced. To begin, I briefly discuss the difficulty of studying predictive algorithms as well as the urgent need to address the inequality produced by these proprietary computer programs. Using critical making as a methodological framework, I argue that Twitter bots are a useful research method for digital rhetoricians studying predictive algorithms. At the heart of this chapter, I describe my personal use of Twitter bots and discuss my difficulty exporting Twitter data for analysis, connecting my experience with the issues of data ownership on the platform. I make no definitive claims about the best or most effective methodology or methods for studying unknown unknowns; instead, I encourage others to adopt a flexible and activist approach to critiquing power structures on social media and other algorithmically mediated digital spaces using my project as inspiration for future work.

Unknown Unknowns of Who to Follow

Twitter's recommendation algorithm—WTF—is a predictive algorithm that curates the types of accounts Twitter users follow and, in turn, information that users are exposed to and speed with which users receive information. The WTF "product," as it is referred to by Twitter, provides highly personalized account recommendations to users with the goal of "maintaining and expanding the active user population" by helping users "discover connections" (Gupta et al.). In 2015, a team of Twitter engineers reported that WTF was directly responsible for more than 500 million new connections each month and produced billions of recommendations a year (Goel et al. 106). At the time, WTF was directly responsible for one-eighth of all connections made on Twitter, not to mention the connections that eventually developed based on initial recommendations (Goel et al. 106). While additional information about WTF's influence has not been released since 2015, WTF is certainly a powerful actor in the Twitterverse.[1]

Although the engineers of WTF have described the general principles of the recommendation algorithm, Twitter users and independent researchers remain in the dark about exactly how the algorithm functions. At the International World Wide Web Conference, the designers responsible for WTF described the recommendation process: using a large-scale snapshot of Twitter's entire network of connections, referred to as an "interest graph," WTF identifies accounts that are "similar to" the user and, from that calculation, accounts the user might be "interested in" (Gupta et al.). Both sets are recommended to the user as potential accounts to follow. And yet, many unknown unknowns remain. For example, how is the "interest graph" developed? How exactly are "similar to" and "interested in" accounts identified? What user data beyond the "interest graph" is used to make recommendations? Is data about user interactions with WTF gathered? Does the algorithm account for the difference in "organic" follows versus "recommended" follows? This lack of information about WTF, as with all other proprietary predictive and recommendation algorithms, results in what Finn Brunton

1. Twitter's WTF is just one of many increasingly influential recommendation algorithms. In their contribution to *The Routledge Handbook of Digital Writing and Rhetoric*, Mihaela Popescu and Lemi Baruh discuss the norming effects of recommendation systems on cultural fields and products. As rhetoric and composition scholars continue to study predictive algorithms, more research focused specifically on the rhetoricity of recommendation systems is needed.

and Helen Nissenbaum refer to as information asymmetry: "when data about us are collected in circumstances we may not understand, for purposes we may not understand, and are used in ways we may not understand" (1-2). Twitter has untold information about our personal lives and lifestyles gathered from sources we do not know about, used in ways we do not know, and results in recommendations that we may not understand.

Studies of both recommendation systems like WTF, it raised red flags about the homogenizing and potentially oppressive effects of the recommendation algorithm. In a study of live recommendation systems, like WTF, researchers found that feedback loops can develop "when a platform attempts to model user behavior without accounting for recommendations" (Chaney et al.). Researchers have also found that recommendation systems using collaborative filtering, as WTF does, are "susceptible to *biases* that may appear in input data," which amplify existing biases and reinforce stereotypes (Tsintzo et al. 1, emphasis in original). Studies of WTF found that the algorithm "disproportionately accelerated the growth of already popular users" likely "altering the diversity of information users consume on the platform" (Su et al.); "further exacerbate[d] the majority-minority gap" by limiting the spread of information (Halberstam and Knight); and created a glass ceiling limiting the visibility of women (Nilizadeh et al.; Zhu et al.) and men of color (Messias et al.). Although there is ample evidence to suggest WTF contributes to inequality on the platform, Twitter's engineers seem unaware or unconcerned. In 2014, a team of Twitter researchers published information about the WTF "interest graph" referring to the graph and their research as "a set of authoritative descriptive statistics" on an active social network (Myers et al. 1). However, at no point in the article do they consider how the implementation of WTF in 2010 affected the structure of Twitter's interest graph and/or contributed to the structure of the network.

Because WTF is part of the larger system of highly influential and inequitable predictive algorithms and Twitter does not seem to hold itself accountable for the effects of their recommendation system, digital rhetoricians, among others, need to continue conducting critical analyses of WTF. And yet, humanities researchers may have less access to the vast resources and technological expertise used to create the big data studies cited above. When I decided to research how WTF may contribute to echo chambers as part of my dissertation, I had significantly limited technological access, support, and know-how. Beyond my personal limitations were the issues of the invisibility of echo chambers and unknown unknowns of predictive algorithms. My ideal goal was to make these invisible and unknowable things somehow tangible for myself as a researcher and for the public that is impacted by recommendation algorithms and echo chambers. While I had a clear vision of my research goals, my approach to completing the project was murkier. As a novice both in terms of research and computer algorithms, I required a methodological framework that supported non-traditional research and was flexible enough to deal with a range of constraints.

A Critical Making Framework

Critical making is a methodological framework for exploring the social aspects of technology through the process of making. As a beginning researcher, critical making appealed to me because of its ad-hoc, do it yourself (DIY) approach that incorporates both academic research and activist work on social media, as evidenced in the collection *DIY Citizenship: Critical Making and Social Media* edited by Matt Ratto and Meagan Boler. As a methodological framework, critical making is especially useful for (1) rendering abstract concepts material through the process of making, (2) humanities and social science scholars studying technology, and (3) process-focused, metacognitive research projects.

Matt Ratto, who popularized the term and founded the Critical Making Lab at the University of Toronto, defines critical making as "materially productive, hands-on work intended to uncover and explore conceptual uncertainties, parse the world in ways that language cannot, and disseminate the results of these explorations through embodied, material forms" ("Textual Doppelgangers" 228). Ratto contends that critical making "frames a need to incorporate technical work alongside critical social analysis and makes a claim that doing so can both extend current scholarly critiques and direct them into society in new ways" ("Textual Doppelgangers" 229). Additionally, critical making focuses on the "constructive process as a site for analysis . . . emphasiz[ing] the shared acts of making rather than the evocative object" ("Critical Making," 253). Ratto and Boler argue that "making as a 'critical' activity . . . provides both the possibility to intervene substantively in systems of authority and power and . . . offers an important site for reflecting on how such power is constituted by infrastructures, institutions, communities, and practices" (1). Critical making, then, provides a flexible methodological framework for activist researchers who want to learn new ways to critique existing power structures through collaborative making.

For researchers who want to experience first-hand the power of algorithms on social media, critical making provides a fruitful framework both for approaching the creation of computer code and the study of unknown unknowns. Rachael Graham Lussos has argued that writing Twitter bots from a critical making framework allows the creator to "experience how the hidden writing of social media technologies—the automated programs that enable (or in some cases, disable) use of those technologies—involves a rhetorical analysis." Although Lussos writes about graduate students, students are not the only academics who can learn about the hidden writing and rules of social media through hands-on experimentation with bots—novice and experienced researchers need to find new ways to engage with and study the predictive algorithms that are increasingly impacting our lives. Because researchers will likely never have access to the proprietary algorithms that they wish to study, the most significant results of our work might be the knowledge we develop through actively engaging with predictive algorithms. Because critical making emphasizes the experiential knowledge that comes from the process of engaging with technologies, digital rhetoricians and researchers interested in intangible and inaccessible algorithms can benefit from adopting such a methodological framework.

Twitter Bots as an Activist Research Method

Twitter bots are small pieces of computer code that interact with Twitter's Application Program Interface (API) to perform certain functions automatically. In this sense, Twitter bots themselves are computer algorithms set to tweet, reply, retweet, or message content based on a predefined set of conditions. Like all algorithms, bots are coded by individuals or groups of individuals. While bots and algorithms run without human intervention and have consequences beyond the intentions of the creator, these programs are not inherently bad, good, or otherwise—they are products of the botmaker's work as well as the culture in which the bots were created. Still, bots are often considered nefarious agents spreading misinformation, propping up authoritarian governments, and generally spamming, annoying, and confusing the Twitter public. These charges are not unfounded: ISIS used social bots to spread radicalism, pro-Russia bots drowned out protest through hashtag manipulation, and social media bots triggered actions in automatic stock market trading systems that resulted in a brief but significant "flash [stock market] crash" (Subrahmanian et al.; Ferrara et al.). While Twitter bots have certainly been used for malicious ends and the effects of bots on the social media ecosystem are complex, Twitter bots can also be tools for activism and social critique. Mark Sample theorizes protest bots or "bots of conviction" as the modern version of a protest song: "a computer program that reveals the injustice and inequality of the world and imagines alternatives." Considering the activist possibilities of Twitter bots, I argue that digital rhetoricians should consider their use for research purposes, while recognizing the complexity of the consequences of bots in digital spaces. Twitter bots can be useful research tools for a range of projects because they require little technical expertise, run automatically, and provide anonymity.

Compared to other automated computer programs, Twitter bots require relatively little technical expertise while giving researchers direct experience creating algorithms. As I will demonstrate later in a description of my collaboration to create Twitter bots, not only is there a large and inviting community of botmakers and enthusiasts who provide online tutorials for creating bots, but researchers can consult with more experienced coders and programmers quite easily.[2] While bots require relatively low technical knowledge, the dividends they pay in conceptual knowledge about social media rhetoric, digital literacy, nonhuman rhetoric, computer programming and automation, among other things,

2. Twitter bots with specific functionalities can also be created using Google spreadsheets that use code developed by more experienced programmers, as documented by Lussos and Holmes and Lussos.

are high. For example, James R. Brown Jr. created a Twitter bot, @yourletterbot, to "grapple" with the realities of the "robot rhetor" (497). Creating the bot helped Brown Jr. conclude that "computation is a rhetorical medium and that software is within the purview of rhetoric" (497). That is, the act of creating bots gives digital rhetoricians invaluable experiential and affectual insights into rhetoric in digital spaces. Making bots and other automated programs actualizes and concretizes abstract and hidden information and mechanisms. Thus, the research produced through the creation of bots can offer robust insights for digital rhetoricians without requiring extensive programming experience.

By their very nature, bots operate automatically, affording researchers and activists a range of benefits from constant data collection to safety from harmful rhetoric. On a very practical level, the automation of bots allows a research project to continue without constant intervention from an individual. For my own research, automation allowed me to gather the content of eight different Twitter feeds twenty-four hours a day, seven days a week. As I will note below, automation also allowed me the ability to gather and archive certain types of data, which prevented me from losing all my research materials after Twitter suspended my accounts. As a form of social critique, automation can create a deluge of counternarrative posts. Steve Holmes and Rachael Graham Lussos argue in an article about bots and #GamerGate[3] that protest bots—such as their own bot @ Dr_Ethics—consistently and constantly injected alternative viewpoints into the one-sided, toxic hashtag stream. Additionally, software engineer Randi Harper used an autoblocking bot to prevent harassment from #GamerGaters before they could even engage with her personal account. Sample considers automation integral to the way bots can protest, as the programs "present society a bill it cannot pay... at the rate of once every two minutes." Automation is certainly powerful and can be used in harmful or annoying ways, but these examples also suggest that automation can be a productive and useful tool for activist researchers to gather data, insert counternarratives and critique into social media platforms, and protect themselves from harassment.

Along the same lines of the protective nature of automation, the anonymity afforded by bots can also protect researchers from online harassment. When developing my research project, one of my advisors' main concerns was my personal safety. Indeed, academics—many of whom are from marginalized and oppressed groups—from across the country and a range of disciplines have been subjected to online harassment from both conservative and liberal extremists (Kamenetz). Through the practice of "doxing," publicly releasing personally identifiable information, online harassment becomes offline threats. Rhetoricians Les Hutchinson and Dana Cloud were both doxxed and targeted for their scholarship. Hutchinson was doxxed and received threats against herself and her family for engag-

3. #GamerGate began in August 2014 as a coordinated harassment campaign against women in tech who spoke out about sexism in the video game industry.

ing in Twitter research for her MA thesis in 2012 (Hutchinson). At RSA 2018, Dana Cloud spoke about the harassment and threats she experienced after being doxxed because of her scholarship. Engaging in research with bots can provide a layer of personal protection and anonymity against online and offline harassment and threats allowing activist researchers to continue their critical-making research.

To help distinguish research bots from bots used to suppress or manipulate users, I suggest that researchers be transparent about their work. I clearly indicated that the accounts were bots through the Twitter handles, names, and bios—all of which also noted that they were part of a research project about echo chambers on social media. I also avoided making the accounts appear human-like by leaving default profile and background images. When my bots followed an account, the profile clearly stated their purpose as research tools. Additionally, as the bot started retweeting content from other users, they could easily remove themselves from the study through blocking the bot. Similarly, researchers who made a Twitter bot to help facilitate social justice organization found that appearing "less human" made the bot more effective in developing connections among users (Savage et al.). Both my bots and Botivist suggest that clearly identifying accounts as bots is not only the most effective method for encouraging user engagement, but also the most ethical.

Critiquing Automation with Automation

While studies suggesting that Who to Follow and other recommendation algorithms have homogenizing and norming effects that are detrimental to minorities further spurred my critical making project, I originally became concerned about echo chambers when I found myself in one. Shortly after joining Twitter in mid-2018, #AsianAugust, which celebrated a historically significant month in Asian American film, began appearing in my feed. While I joined Asian American actors, filmmakers, and fans on Twitter in cheering over the release of *Crazy Rich Asians*, *To All the Boys I've Loved Before*, and *Searching*, my offline friends and colleagues seemed to know little about #AsianAugust. Being part of the #AsianAugust echo chamber disconcerted me on two fronts. First, I was surprised at how distorted my perspective on #AsianAugust was. Because the topic was so popular on my Twitter feed, I assumed it was popular on everyone else's—the issue had been amplified within my Twitter echo chamber. Second, I was dismayed that #AsianAugust was not gaining more widespread attention outside the Asian American community. The hashtag was being suppressed or filtered out of the content feeds of others. When I began designing research projects for my dissertation, I remembered my frustration, alarm, and disbelief about the (lack of) circulation of #AsianAugust.

Although the project was a solo endeavor, I drew on the expertise and guidance of a range of collaborators—online forums, open-source code and applications, online videos, national coding organizations, computer programmers, and a fellow graduate student—to execute this critical-making study. When I conceived of the retweeting bots, I only had a surface-level knowledge of the capability of bots and no experience coding. I found bot enthusiast Stefan Bohacek's Botwiki forum and began watching Daniel Shiffman's "The Code Train" YouTube tutorial series on Twitter bots. I attended a Women Who Code event hosted by the Dallas-Fort Worth chapter where I worked with a local computer programmer who specialized in Node.js. We collaborated not only on writing the Twitter bot code, but also thinking through how the bot would perform the functions I needed to make individual content feeds visible to the public. To set up automated retweeting as well as archive data on a database service, I worked with a fellow graduate student, Sean McCullough, who had more experience in programming. I list the steps for creating a bot not only to document the process for myself and others, but to point to the many types of collaboration I used for this critical making project. While working with a computer programmer and graduate student are traditional modes of collaboration, I also collaborated using open-source code and applications, online tutorials, and virtual forums. These digitally mediated modes of collaboration significantly contributed to my project, just in-person collaboration did.

Through these collaborations, I designed a bot that would gather the last 200 unique tweets that appeared on the account's content feed and retweet every other tweet at a thirty-minute interval. Retweeting at these intervals allowed the bot to consistently retweet content, but not tweet beyond the daily and hourly limits imposed by Twitter ("About Twitter Limits"). When completed, the bots performed three main functions each time the program ran:

1. Gather the last 200 tweets that appeared on the account's content feed
2. Retweet every other gathered tweet to the account's timeline
3. Send retweet data (Twitter handle, date and time, and text-based content, among other unique identifiers) to a database for archiving

From July 1 to November 5, 2019, bots retweeted the content from eight different Twitter feeds and archived data about the tweets.

The process of designing and coding the bots taught me about how algorithms must be intentionally designed, but also how quickly and easily algorithms can be created. Leigh Gruwell has challenged digital rhetoricians to work "within the confines of the platforms they study" to "take advantage of each space's unique affordances" (Gruwell). Twitter created the "retweet" function to encourage user engagement through the recirculation of content. I used this feature of Twitter's architecture because the retweeting function (1) made the content of personalized Twitter feeds publicly visible and (2) allowed me to create an archive of each account's content feed. Without exploiting the retweeting function, I would have been unable to make the content visible or analyzable. My main takeaway from coding the retweeting bot was that any algorithm designer needs to be highly attentive to the unexpected outcomes of their work and consult with many different stakeholders to avoid accidentally and later, automatically harming others.

The process of following accounts using Twitter's WTF recommendation algorithm gave me insight into how individuals experience the creation of a social media echo chamber and how influential WTF is in creating personalized feeds. Because recommendation systems "are a paradigmatic example of the interaction between humans and algorithms in the cultural arena," not only do echo chambers reflect user and algorithm bias, but the Twitter feeds I created needed to account for both (Bressan et al. 745). To begin, I identified highly influential accounts within a given conversation. These accounts would be considered "interested in" accounts that are highly vocal in a particular discussion. After following accounts from highly influential and prominent members of a community, I transitioned to use WTF recommendations, allowing the algorithm to take over the following process. During this phase of following, WTF rapidly served up recommendations, building an archive of data on my preferences and further pushing each Twitter account into groups. I also followed accounts that appeared within the content feed. When an individual reposts content on their timeline, it often appears in the timelines of their followers. While these secondary account follows are not directly coordinated by WTF, they are still influenced by the recommendation algorithm.

Establishing methods for coding Twitter bots and following Twitter accounts took substantial time and thought, but the creation of the personalized content feeds that could be considered echo chambers took very little time. I initially planned on gathering data about the echo chambers over several months, assuming that it would take time for the divisions to appear across the eight different Twitter accounts. However, I was surprised to see just how quickly personalized Twitter feeds that highlighted specific worldviews developed. For each Twitter feed, I followed approximately 100 accounts in less than thirty minutes. With WTF serving up hundreds of new recommendations every second, each account rapidly developed a distinct network of "friendships," the term Twitter uses for accounts a user follows. There was only a small fraction—5.5 percent—of overlap across the eight accounts' "friendships." Although I could only mimic the perspectives of a particular group, the Who to Follow algorithm quickly and efficiently created personalized Twitter feeds reflecting those perspectives.

This small study suggests—and Twitter appears to confirm by terminating my activity—that the platform itself is designed to manipulate and disrupt the exchange of information through the creation of personalized feeds. Twitter relies heavily on user interactions, and feeds that support an individual's pre-existing views are one of the best ways to increase engagement. Feeding people what they want to "like" and recirculate makes perfect sense from a user engagement-focused perspective. Although I had read plenty of research about how WTF contributes to echo chambers and increases inequality, the process of creating personalized feeds representing specific viewpoints made the abstract concept far more concrete and disturbing. I understood from hands-on experience just how quickly people are sorted, categorized, and pigeonholed by the WTF recommendation algorithm.

Data Problems

During the conceptual phases of my critical making study, one of the main concerns was the sheer amount of data that the retweeting bots would generate. With eight accounts potentially tweeting a combined 38,400 tweets a day, the amount of raw data could become unmanageable quickly. However, obtaining usable and analyzable data was the most difficult and unexpected obstacle I encountered during this research project. When all eight Twitter accounts were suspended on November 5, 2019, I immediately lost access to each account's timelines as well as information about the followed and follower accounts. Even before I was shut out of Twitter, I had difficulty gathering data that included media-rich content. The goal of this project was to understand how individuals experience personalized Twitter feeds, so I wanted to qualitatively analyze tweets as they would appear in a Twitter feed, not as decontextualized strings of text-based data best suited for corpus analysis. Gathering qualitatively analyzable data proved the most significant and persistent challenge of this project and speaks to larger issues of data ownership on social media.

Despite my best efforts and consultation with others, I found no method of data exporting that could provide me with media-rich data that was unfiltered by Twitter. ATLAS.ti, qualitative analysis software, was the only platform I found that offered media-rich content, displaying tweets in a way like how they would appear on Twitter. However, the software offers researchers limited ability to know exactly how the imported tweets were chosen; ATLAS.ti data is a selection of 100 "recent" tweets that are mediated by Twitter. ATLAS.ti acknowledges the limitations of their Twitter exports in the user manual: "Note that you only will be able to import tweets from the last week. Further, as the final selection is done by Twitter and not within our control, queries at different times, or on different computers may result in different tweets" (110). To my knowledge, ATLAS.ti's interface is the only way to export tweets easily and quickly in a way that replicates the experience of individual users. Twitter's insistence on filtering the exporting of data replicates the content feed personalization on the platform itself. It remains an unknown unknown how and why any tweet appears on a users' content feed and the same goes for exported tweets on ATLAS.ti.

My difficulty accessing data from Twitter is in line with the platform's Terms of Service that give very little power to users and virtually unlimited power to the tech company. By simply submitting, posting, or displaying content on Twitter, users "grant [the platform] a worldwide, non-exclusive, royalty-free license (with the right to sublicense) to use, copy, reproduce, process, adapt, modify, publish, transmit, display and distribute such Content in any and all media or distribution methods now known or later developed" ("Twitter Terms of Service"). Twitter also retains the right to "suspend or terminate your account or cease providing you with all or part of the Services at any time for any or no reason" ("Twitter Terms of Service"). In other words, Twitter takes an authoritarian approach on its own platform—the company not only wields absolute authority over their proprietary algorithms, but also over content produced by individuals using the platform.

While researchers and individual Twitter users have trouble accessing their own data and knowing how their data is used by the platform, research using bots may raise concerns about contaminating data on social interactions and network connections on Twitter. Because Twitter's recommendation algorithm uses a snapshot of the existing network architecture to calculate recommendations, experimenting with WTF could be contaminating the "interest graph." Even worse, by allowing WTF to prompt me to create echo chambers, some might argue that my project amplifies and reinforces the homogenizing effects of the recommendation algorithm.

These are valid concerns about my project and others that attempt to experiment with live social media platforms. However, I argue that small studies such as mine have very little influence on the overall social structure of social media platforms and, even more importantly, data about social interactions on social media is already compromised by predictive algorithms that feed on and exacerbate implicit bias. Brunton and Nissenbaum argue that "data pollution is unethical only when the integrity of the data flow or data set in question is ethically required" (69). Twitter's use of the WTF recommendation system without accounting for the homogenizing effects of the algorithm already compromises the data set. Additionally, because I studied unknown unknowns from a weak position in the information asymmetry, I would argue that my methods are justifiable. Nevertheless, issues surrounding proprietary algorithms and ownership, management, and contamination of the user data gathered by and fueling these algorithms remains a thorny issue that will not be resolved if tech companies keep the public in the dark.

Conclusion

Studying unknown unknowns can be incredibly frustrating. I hit roadblock after roadblock trying to export data and reinstate suspended Twitter accounts. At the time of writing, my personal Twitter account remains suspended, and Twitter has refused to provide additional information about why I was suspended or how I might be able to return to the platform. With only a list of "friendships" for each account and a handful of tweets filtered by Twitter, I had to work with limited data. However, these frustrations and setbacks have helped me further understand the power dynamics of social media platforms. Twitter quite clearly exerted its power over me as an individual and activist researcher. While I was still able to make insights about how WTF contributes to the creation of echo chambers, experiencing the difficulty of studying unknown unknowns and feeling the full force of Twitter's authority has been the most significant learning experience from this study. Some of the tangible products of my research may have been erased, but the deep experiential knowledge that I developed during this critical making project remains.

For activist researchers interested in adopting a critical making methodological framework, studying unknown unknowns, using automated bots for research, or experimenting on a live social media platform, I offer the following suggestions for designing a research project:

- Look to other academic disciplines and activists for new research methods
- Consider the affordances of your chosen platform of study and incorporate the platform into your critique
- Collaborate in a range of modes when developing your research project
- Document and reflect on the creation process and prototyping phase of any objects or artifacts that you make
- Export data as frequently as possible to multiple platforms, but know that you may need to adjust your research goals and/or results based on the ability to retain analyzable data

If social media platforms like Twitter had their way, independent researchers would never gain access to information about echo chambers or predictive algorithms. The data gathered about individuals and groups is the currency of the internet and any technologies, such as recommendation systems, that increase user engagement or data collection are highly valuable to social media corporations. It is precisely because these unknown unknowns are so securely guarded, profitable, and influential that digital rhetoricians and other academics need to conduct publicly accessible scholarship. Thus, adopting non-traditional, activist research methodologies is imperative to increase public and scholarly knowledge about predictive algorithms, content circulation, and echo chambers.

Works Cited

About Twitter Limits. Twitter, https://help.twitter.com/en/rules-and-policies/twitter-limits. Accessed 22 July 2020.

ATLAS.ti 8 Mac User Manual. ATLAS.ti Scientific Software Development, Berlin, 2019.

Beck, Estee. "Implications of Persuasive Computer Algorithms." *The Routledge Handbook of Digital Writing and Rhetoric*. edited by Jonathon Alexander and Jacqueline Rhodes, Routledge, 2018, pp. 291-300.

Bressan, Marco, et al. "The Limits of Popularity-Based Recommendations, and the Role of Social Ties." KDD Conference, 13-17 Aug. 2016, San Francisco, California.

Brown Jr., James J., "The Machine That Therefore I Am." *Philosophy and Rhetoric*, vol. 47, no. 4, 2014, pp. 494-512.

Brunton, Finn, and Helen Fay Nissenbaum. *Obfuscation: A User's Guide for Privacy and Protest*. MIT Press, 2015.

Chaney, Allison J.B., Brandon M. Stewart, and Barbara E. Engelhardt. "How Algorithmic Confounding in Recommendation Systems Increases Homogeneity and Decreases Utility." ACM Conference on Recommender Systems (RecSys) held at Vancouver, Canada, 2018.

Cinelli, Matteo, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. "Echo Chambers on Social Media: A Comparative Study." *ArXiv:2004.09603 [Physics]*, Apr. 2020, http://arxiv.org/abs/2004.09603.

Cloud, Dana. "Fighting Back Against the Alt-Right's Higher Education Agenda." *18th Biennial Conference and 50th Anniversary Celebration of the Rhetoric Society of America* held in Minneapolis, Minnesota, 2018.

"The Coding Train." YouTube, https://www.youtube.com/channel/UCvjgXvBlbQi-ydffZU7m1_aw. Accessed 22 July 2020.

Del Vicario, Michela, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. "The Spreading of Misinformation Online." *Proceedings of the National Academy of Sciences*, vol. 113, no. 3, 2017, pp. 554-59.

@Dr_EthicsGate. May 2017, https://twitter.com/Dr_EthicsGate.

Ferrara, Emilio, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. "The Rise of Social Bots." *Communications of the ACM*, vol. 59, no. 7, July 2014.

Gallagher, John R. "Writing for Algorithmic Audiences." *Computers and Composition*, vol. 45, 2017, pp. 25-35.

Garimella, Kiran, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. "The Effect of Collective Attention on Controversial Debates on Social Media." *WebSci '17: 9th International ACM Web Science Conference*, 2017, pp. 43-47.

Goel, Ashish, Pankaj Gupta, John Sirois, Dong Wang, Aneesh Sharma, and Siva Gurumurthy. "The Who-To-Follow System at Twitter: Strategy, Algorithms, and Revenue Impact." *INFORMS Journal on Applied Analytics*, vol. 45, no. 1, 2015, pp. 98-107.

Gruwell, Leigh. "Constructing Research, Constructing the Platform: Algorithms and the Rhetoricity of Social Media Research" *Present Tense*, vol. 6, no. 3, 23 Jan. 2018, http://www.presenttensejournal.org/volume-6/constructing-research-constructing-the-platform-algorithms-and-the-rhetoricity-of-social-media-research/.

Gupta, Pankaj, Ashish Goel, Jimmy Lin, Aneesh Sharma, Dong Wang, and Reza Zadeh. "WTF: The Who to Follow Service at Twitter." International World Wide Web Conference Committee, 13-17 May 2013, Rio de Janeiro, Brazil.

Halberstam, Yosh, and Brian Knight. "Homophily, Group Size, and the Diffusion of Political Information in Social Networks: Evidence from Twitter." *Journal of Public Economics*, vol. 143, 2016, pp. 73-88.

Holmes, Steve, and Rachael Graham Lussos. "Cultivating Metanoia in Twitter Publics: Analyzing and Producing Bots of Protest in the #GamerGate Controversy." *Computers and Composition*, vol. 48, June 2018, pp. 118–38.

Hutchinson, Les. “Wielding Power and Doxing Data: How Personal Information Regulates and Controls our Online Selves.” *The Routledge Handbook of Digital Writing and Rhetoric*. edited by Jonathan Alexander and Jacqueline Rhodes, Routledge, 2018, pp. 303-16.

Kamenetz, Anya. “Professors Are Targets in Online Culture Wars; Some Fight Back.” *NPR.Org*, 4 Apr. 2018, https://www.npr.org/sections/ed/2018/04/04/590928008/professor-harassment.

Lussos, Rachael Graham. “Twitter Bots as Digital Writing Assignments.” *PraxisWiki* on *Kairos: Rhetoric, Technology, and Pedagogy*, vol. 22, no. 2, 19 Dec. 2017, http://praxis.technorhetoric.net/tiki-index.php?page=PraxisWiki%3A_%3ATwitter-Bots#References.

Messias, Johnnatan, Pantelis Vikatos, and Fabrício Benevenuto. “White, Man, and Highly Followed: Gender and Race Inequalities in Twitter.” International Conference on Web Intelligence held in Leipzig Germany, 2017.

Mitchell, Amy, Mark Jurkowitz, J. Baxter Oliphant, and Elisa Shearer. “Three Months In, Many Americans See Exaggeration, Conspiracy Theories and Partisanship in COVID-19 News.” *Pew Research Center’s Journalism Project*, 29 June 2020, https://www.journalism.org/2020/06/29/three-months-in-many-americans-see-exaggeration-conspiracy-theories-and-partisanship-in-covid-19-news/.

Myers, Seth A., Aneesh Sharma, Pankaj Gupta, and Jimmy Lin. “Information Network or Social Network? The Structure of the Twitter Follow Graph.” International World Wide Web Conference Committee held in Seoul South Korea, 1-11 Apr. 2014.

New York Supreme Court, Appellate Term—First Department. *The People of the State of New York v. Malcolm Harris and Twitter, Inc*. American Civil Liberties Union, https://www.aclu.org/files/assets/brief-harris-appellateterm.pdf. Accessed 20 July 2020.

Nilizadeh, Shirin, Anne Groggel, Peter Lista, Srijita Das, Yong-Yeol Ahn, Apu Kapadia, and Fabio Rojas. “Twitter’s Glass Ceiling: The Effect of Perceived Gender on Online Visibility.” AAAI Conference on Web and Social Media held in Cologne Germany, 2016.

Popescu, Mihaela, and Lemi Baruh. “Privacy as Cultural Choice and Resistance in the Age of Recommender Systems.” *The Routledge Handbook of Digital Writing and Rhetoric*. edited by Jonathan Alexander and Jacqueline Rhodes, Routledge, 2018, pp. 280-90.

Ratto, Matt. “Critical Making: Conceptual and Material Studies in Technology and Social Life.” *The Information Society*, vol. 27, 2011, pp. 252-60.

---. “Textual Doppelgangers: Critical Issues in the Study of Technology.” *DIY: Citizenship: Critical Making and Social Media*. edited by Matt Ratto and Megan Boler, MIT Press, 2014, pp. 227-37.

Ratto, Matt and Megan Boler. “Introduction.” *DIY: Citizenship: Critical Making and Social Media*. edited by Matt Ratto and Megan Boler. MIT Press, 2014, pp. 1-22.

“Republicans, Democrats Move Even Further Apart in Coronavirus Concerns.” *Pew Research Center - U.S. Politics and Policy*, 25 June 2020, https://www.people-press.org/2020/06/25/republicans-democrats-move-even-further-apart-in-coronavirus-concerns/.

Sample, Mark. "A Protest Bot Is a Bot so Specific You Can't Mistake It for Bullshit." *Medium*, 4 Oct. 2015, https://medium.com/@samplereality/a-protest-bot-is-a-bot-so-specific-you-cant-mistake-it-for-bullshit-90fe10b7fbaa.

Savage, Saiph, et al. "Botivist: Calling Volunteers to Action Using Online Bots." *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing - CSCW '16*, ACM Press, 2016, pp. 811–20.

Su, Jessica, Aneesh Sharma, and Sharad Goel. "The Effects of Recommendations on Network Structure." World Wide Web Conference 2016 held in Montreal, Canada, 11-15 Apr. 2016.

Subrahmanian, V.S., et al. "The DARPA Twitter Bot Challenge." *Computer*, vol. 49, no. 6, June 2016, pp. 38-46.

Suciu, Peter. "More Americans Are Getting Their News from Social Media." Forbes, 11 Oct. 2019, https://www.forbes.com/sites/petersuciu/2019/10/11/more-americans-are-getting-their-news-from-social-media/.

Sunstein, Cass R. *Republic.Com*. Princeton UP, 2002.

Twitter Terms of Service. Twitter, 18 June 2020, https://twitter.com/en/tos#us. Accessed 14 July 2020.

Vee, Annette. "Understanding Computer Programming as a Literacy." *Literacy in Composition Studies*, vol. 1, no. 2, 2013, pp. 42-64.

@yourletterbot. Aug. 2013, https://twitter.com/yourletterbot. Accessed 5 July 2020.

Zhu, Jane M., Arthur P. Pelullo, Sayed Hassan, Raina M. Merchant, and Rachel M. Werner. "Gender Differences in Twitter Use and Influence Among Health Policy and Health Services Researchers." *JAMA Internal Medicine*, vol. 179, no. 12, Dec. 2019, pp. 1726-29.