



Selected Papers of Internet Research 16:  
The 16<sup>th</sup> Annual Meeting of the  
Association of Internet Researchers  
Phoenix, AZ, USA / 21-24 October 2015

## **BOT-BASED COLLECTIVE BLOCKLISTS IN TWITTER: THE COUNTERPUBLIC MODERATION OF A PRIVATELY-OWNED NETWORKED PUBLIC SPACE**

R. Stuart Geiger

University of California, Berkeley

### **Abstract**

In Twitter, many people are facing increasing harassment and abuse from others, particularly from individuals associated with "GamerGate"— a self-described 'social movement' viciously opposing feminist video game developers and media critics. While Twitter supports user-to-user 'blocking' (anyone can direct Twitter to hide posts or messages from a particular account), targets of GamerGate-associated harassment often describe individually blocking harassers as a Sisyphean task. In response, some are using collective blocklists, in which a group curates a list of accounts which they have identified as harassers. Any account added to a blocklist is automatically made invisible for all of the blocklist's subscribers. Most notably, this feature is not built into the Twitter platform and was not designed, developed, or officially supported by Twitter, Inc. Instead, collective blocklists are made possible through automated software agents (or bots) which are developed and operated by independent groups of volunteers.

This paper reports findings from an ethnography of infrastructure, investigating the development and deployment of these bot-based collective blocklists (or blockbots) in Twitter. I show how the designs (and re-designs) of blockbots are bound up in competing ideas and imaginaries about what it means for counterpublic groups to moderate a privately-owned networked public space. Blockbots are a mode of algorithmic filtering which reconfigures affordances of a networked public, but with key differences from the algorithmic filters like in Facebook's News Feed. Blockbots make responding to harassment a more visible and communal practice, as well as involve imagining alternative policies and procedures for moderating content online.

Suggested Citation (APA): Geiger, R.S. (2015, October 21-24). *Bot-based collective blocklists in Twitter: The counterpublic moderation of a privately-owned networked public space*. Paper presented at Internet Research 16: The 16<sup>th</sup> Annual Meeting of the Association of Internet Researchers. Phoenix, AZ, USA: AoIR. Retrieved from <http://spir.aoir.org>.

## Research questions and methodology

This paper reports from a study of bot-based collective blocklists (or blockbots) in Twitter, methodologically drawing from techniques and tactics in Star's ethnography of infrastructure (Star 1999). Twitter is an infrastructure supporting participation in a networked public (boyd 2010), but blockbots are also infrastructures supporting a set of practices in which counterpublic groups (Fraser 1990) work to moderate a privately-owned public space. I focus on how these infrastructures are not static, but rather dynamic and relational, "emerg[ing] for people in practice, connected to activities and structures" (Bowker et al. 2010: 99). I iteratively and inductively used many methods – interviews with blockbot developers and subscribers, observations of routine and exceptional activity, archival and historical methods, and systems analysis and software studies methods. My overarching research question was how blockbots make visible and contest particular assumptions about how Twitter operates as a networked public.

Within this broader research question, I had two specific sub-questions: First, how are blockbots currently operating in Twitter – what are the average, everyday practices in their operation? Second, how has each specific blockbot developed and evolved over time, due to both internal and external factors? I focused on the organizational structure behind specific blockbots and sought to understand the different norms and discourses used by those in and around each blockbot. I also sought moments of controversy and breakdown to make structures, norms, discourses, and invisible work (Star and Strauss 1999) visible and comparable. While I focused on blockbot developers and subscribers, these controversies also led me to observing activity and analyzing publicly-accessible discourse from a wide variety of social actors in various environments, including staff at Twitter, Inc., anti-harassment activists, commentators in social and mainstream media, and GamerGate-aligned groups.

## Findings and implications

*Blockbots are ways in which people who do not have access to server-side code seek to change or contest the default affordances of platforms*

Lessig's 'code is law' (1999) maxim is a powerful metaphor to conceptualize software in society: it casts organizations like Twitter, Inc. as governments, whose developers have the exclusive authority to legislate from the command line. Typically, if people disagree with how Twitter is designed or administered – such as how many people believe Twitter, Inc. is not doing enough to stop harassment – they can express themselves by stating their objections or by leaving the platform. In order to change the software-based affordances of a platform like Twitter, generally, people must convince Twitter, Inc. to change the code-as-law these server-sovereigns have written and implemented.

Bot-based collective blocklists in Twitter involve a third approach: they are "bespoke code" (Geiger 2014) in which people who do not have access to modify a platform's server-side codebase are nevertheless able to change how the platform operates. And as boyd (2010) argues, differently-programmed networked public spaces can have wildly different affordances in terms of how persistent, replicable, scalable, or searchable

content is. Blockbots are explicitly and intentionally designed, developed, and deployed to change these affordances. Most are implemented because of a perceived failure by Twitter, Inc. to adequately deal with harassment. Some who develop and operate Twitter blockbots are also actively seeking to alter code-dependent affordances on the 'server side' – hoping for a day when their blockbots are obsolete.

*Blockbots make responding to harassment a more visible and communal practice*

While Twitter does have a 'blocking' feature which lets users hide accounts, blockbots turn this activity from an individual to a collective practice. As Crawford and Gillespie (2014) argue, mechanisms of 'flagging' can shift responsibility for policing content from the companies operating platforms to the people using them. This devolution in responsibility can be celebrated as giving more agency to people who use such platforms, but it also shifts the burden of conducting such moderation work to individuals, which can involve a substantial amount of potentially-traumatic labor.

By default in Twitter, harassment can be (and is) easily coordinated by like-minded groups, who distribute labor in ways that are both efficient and visible to each other, particularly with coordinated harassment campaigns like Gamergate (Chess and Shaw 2015; Heron, Belford, and Goker 2014). As responding to harassment involves an individual, private client-to-server directive, this work is far less visible and harder to coordinate with one's peers and allies. Bot-based collective blocklists make the affective labor of responding to coordinated harassment not only more efficient, but also more visible and communal. Communities have formed around many blockbots, whose developers, moderators, and subscribers give each other a wide variety of logistical, emotional, and institutional support – in addition to the average, everyday work involving in collectively curating a blocklist.

*Blockbots imagine alternative policies and procedures for moderating content online*

In addition to implementing alternative code-based affordances, blockbot groups are exploring and developing alternative policies and procedures for moderating content online. They engaging in the same kinds of tasks as Twitter's Trust and Safety team, but with many differences. And as many of these groups have grown from 'teams of one' to dozens of members, many have developed increasingly-formalized procedures for reviewing content, as well as developing their own norms and discourses. Yet even organizations that share similar tasks and goals often have different procedures and norms. And like any group, they often debate these issues both among themselves and with those outside of the organization.

*Blockbots are a mode of algorithmic filtering, but invert several default assumptions*

While there is much debate on the algorithmic filtering of posts in the 'news feeds' or 'timelines' of social networking sites (Bucher 2012), much of the public and academic discourse on this issue assumes that such algorithmic filtering operates as it currently does in Facebook: a single opaque 'black box', developed internally by the company that owns and operates the site, which is enabled by default and difficult to opt-out of. Yet bot-based collective blocklists are a form of algorithmic filtering that generally do not operate

according to any of these assumptions, providing a compelling alternative when considering the socio-technical governance of networked publics.

## References

- Bowker, Geoffrey C., Karen Baker, Florence Millerand, and David Ribes. 2010. "Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment." In *International Handbook of Internet Research*, 97–117. doi:10.1007/978-1-4020-9789-8\_5.
- boyd, d. 2010. "Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications." In *Networked Self: Identity, Community, and Culture on Social Network Sites*, edited by Zizi Papacharissi, 39–58. <http://www.danah.org/papers/2010/SNSasNetworkedPublics.pdf>.
- Bucher, T. 2012. "Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook." *New Media & Society* 14 (7): 1164–80. doi:10.1177/1461444812440159.
- Chess, Shira, and Adrienne Shaw. 2015. "A Conspiracy of Fishes, Or, How We Learned to Stop Worrying About #GamerGate and Embrace Hegemonic Masculinity." *Journal of Broadcasting & Electronic Media* 59 (March 2015): 37–41. doi:10.1080/08838151.2014.999917.
- Crawford, Kate, and Tarleton Gillespie. 2014. "What Is a Flag for? Social Media Reporting Tools and the Vocabulary of Complaint." *New Media & Society* 2. doi:10.1177/1461444814543163.
- Fraser, Nancy. 1990. "Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy." *Social Text* 26: 56–80. doi:10.2307/466240.
- Geiger, R. Stuart. 2014. "Bots, Bespoke, Code and the Materiality of Software Platforms." *Information, Communication & Society* 17 (3). Routledge: 342–56. doi:10.1080/1369118X.2013.873069.
- Heron, Michael James, Pauline Belford, and Ayse Goker. 2014. "Sexism in the Circuitry." *ACM SIGCAS Computers and Society* 44 (4). ACM: 18–29. doi:10.1145/2695577.2695582.
- Lessig, Lawrence. 1999. *Code and Other Laws of Cyberspace*. New York: Basic Books.
- Star, Susan Leigh. 1999. "The Ethnography of Infrastructure." *American Behavioral Scientist* 43 (3): 377–91.
- Star, Susan Leigh, and Anselm Strauss. 1999. "Layers of Silence , Arenas of Voice : The Ecology of Visible and Invisible Work." *Computer Supported Cooperative Work* 8: 9–30. doi:10.1023/A:1008651105359.