Accepted manuscript version:
Lally, N. 2016. Crowdsourced surveillance and networked data. *Security Dialogue.*
http://sdi.sagepub.com/content/early/2016/09/03/0967010616664459

Crowdsourced surveillance and networked data

Nick Lally
University of Wisconsin–Madison

Abstract
Possibilities for crowdsourced surveillance have expanded in recent years as data uploaded to social networks can be mined, distributed, assembled, mapped, and analyzed by anyone with an uncensored internet connection. These data points are necessarily fragmented and partial, open to interpretation, and rely on algorithms for retrieval and sorting. Yet, despite these limitations, they have been used to produce complex representations of space, subjects, and power relations as internet users attempt to reconstruct and investigate events while they are developing. In this article, I consider one case of crowdsourced surveillance that emerged following the detonation of two bombs at the 2013 Boston Marathon. I focus on the actions of a particular forum on reddit.com, which would exert a significant influence on the events as they unfolded. This study describes how algorithmic affordances, internet cultures, surveillance imaginaries, and visual epistemologies contributed to the structuring of thought, action, and subjectivity in the moment of the event. I use this case study as a way to examine moments of entangled political complicity and resistance, highlighting the ways that particular surveillance practices are deployed and feed back into the event amidst its unfolding.

Keywords
Algorithms, crowdsourcing, data, social networks, surveillance

Introduction

On the afternoon of April 15th, 2013, two bombs exploded near the crowded finish line of the Boston Marathon, killing three people and injuring more than 250. It would be three days before the FBI would publicly identify Dzhokhar and Tamerlan Tsarnaev as suspects in the case.

Meanwhile, the FBI made public calls for the community to assist the investigation by submitting information related to the bombings. "We are particularly interested in reviewing video footage captured by bystanders with cell phones or personal cameras near either of the blasts," wrote Attorney General Tim Holder in an FBI press release the day following the attack (FBI, 2013a). Special Agent in Charge Richard DesLauriers made a similar request in a press conference on the same day, asking the public to alert the FBI if they had noticed "Someone who appeared to be carrying an unusually heavy, dark-colored bag yesterday around the time of the blasts and in the vicinity of the blasts" (FBI, 2013b). The suggestion that the thousands of images and videos from bystanders might contain the clues needed to find the suspect with the heavy, dark bag was enough to set thousands of untrained internet detectives into motion. These internet detectives, working outside of the framework of the FBI call, would exert a significant influence on the events that transpired following the bombing. The FBI, who conjured up this internet surveillance machine, would soon find itself unable to control the forces it had summoned.

The initial call from the FBI, it should be noted, was not anomalous in regards to the current logics of American national security. The Department of Homeland Security, for example, encourages people to report unusual or suspicious activity through their "If You See Something, Say Something" campaign, which was developed in response to the perceived risk of terrorism. The campaign's website claims that citizen vigilance plays "a critical role in keeping our nation safe" (United States Department of Homeland Security, n.d.). This logic of distributed surveillance is certainly not new – we see familiar logics running through neighborhood watch programs, road signs asking to report drunk drivers, McCarthy era witch hunts, Rumor Control Centers of the 1960s (Young et al., 2014) and even, as some scholars have argued, in 5th century practices of governance (Reeves, 2012). But recent efforts to encourage

citizen vigilance have developed alongside the proliferation of personal technologies, cell phones and digital cameras in particular, that enable the recording and sharing of vast amounts of data. It is not uncommon, then, for law enforcement institutions to 'crowdsource' parts of an investigation by issuing an open call for information in an effort to gather some of these data.

Coined by Jeff Howe in a 2006 *Wired* article, the term 'crowdsource' is an amalgamation of 'crowd' and 'outsource.' It refers to a "distributed problem-solving and production model that leverages the collective intelligence of online communities," where the crowd is made up of the online community (Brabham, 2013: xix). Typically, a call is put out by an organization requesting participants to engage in small, often creative tasks to achieve an organizational goal. Participants who take up the call, whether enticed by money, belief in the project, personal interest, or some other incentive, perform tasks that fall within the framework of the organization. In this way, crowdsourcing mixes bottom-up creativity with top-down managerial organization (Brabham, 2013: 4). While the model is often applied to business projects (Howe, 2009), it has also been used effectively in science (Bhardwaj, 2014), policy (Brabham, 2013: 34), and policing contexts (Schneider and Trottier, 2012). Crowdsourcing takes as its premise the idea that harnessing the intelligence of groups of people, under the correct conditions, can produce knowledge that exceeds the possibilities of the best-trained individuals (Surowiecki, 2004). In the case of the FBI call, little creative work was required by the crowd. Trained professionals would ultimately be in charge of making sense of the contributed data.

While the FBI received a wealth of information following its call, images and photos from bystanders also quickly found their way to Twitter, YouTube, and other social media sites. Many were then picked up by internet and television news media in the hours following the bombing and included in reports. Almost immediately, the state, media, witnesses, and media consumers became connected through complex communication networks as data to reconstruct an

understanding of the event were produced, distributed, analyzed, shared, and consumed. Additionally, the circulation of these data on social media afforded new possibilities for participation in the event. Data uploaded to social networks could now be mined, assembled, mapped, and analyzed by anyone with an uncensored internet connection. These data points are necessarily fragmented and partial, open to interpretation, and limited to a small sliver of public data that are part of much larger corporate databases. Yet, despite these limitations, they were used to produce complex representations of space, subjects, and power relations as internet users attempted to reconstruct and investigate the event amidst its unfolding.

Participation by a growing group of internet users quickly exceeded the parameters set by the FBI. Internet users from around the world, many of them convening on online message boards 4chan and Reddit, began collecting and analyzing hundreds of images uploaded to social media in an effort to reconstruct the event and find the bombers. In this article, I focus on the /r/findbostonbombers subreddit[1]—a forum on reddit.com made for users to "compile, analyze, and discuss images, links, and thoughts about the Boston Bombing." The investigations on the subreddit would exert a significant impact on the unfolding event as these online discussions filtered into the broader world in various ways.

The subreddit created its own structure for crowdsourcing the investigation, which differed significantly from the FBI call. Instead of a crowdsourcing structure organized by the state, the subreddit relied on the technological affordances of the Reddit platform, the gatekeeping role of a moderator, and the interactions of the crowd that formed around the forum. The group of users, who call themselves 'redditors', that coalesced within this forum formed what danah boyd (2011:39) has termed a 'networked public.' This public, she explains, is an imagined community within a space structured by technology. This structure "introduces distinct affordances that shape how people engage with the environment." This is not to say that

these structures determine practice, but rather, that users must contend with them as they participate in the platform (boyd, 2011: 55). Structures shape and are shaped by a variety of forces—economic, social, cultural, and political—that intersect with their production and use (Langlois and Elmer, 2013: 5).

Central to this technological structuring are algorithms—the coded instructions that determine how computational processes function. Algorithms establish a framework that contributes to shaping the possibilities for action as they organize the vast surveillance assemblage that became visible in Boston. They act in situated contexts with contingent and unpredictable outcomes (Kitchin, 2014: 22), influencing how we perceive and think about the world (Gillespie, 2014: 183) "Algorithms act, but they do so as part of an ill-defined network of action upon actions," (Goffey, 2008: 19) making them highly relational and difficult to grasp. Since these outcomes are unpredictable and situated, this study focuses on the single case of the Boston bombing to understand how algorithmic affordances contribute to the structuring of thought, action, and subjectivity in the moment of the event. And so we begin the in the middle of the event, in the moments following the two blasts, as internet users began tapping into online data streams, revealing new possibilities for crowdsourced surveillance.

The Immediate Aftermath

> "Man, I'm never going to a never [sic] public event out of fear that I may glance away for a second and become a suspect." —redditor

Following the bombings, users on Reddit, the 32nd most popular website in the world and 9th most popular in the United States at the time of writing (Alexa, 2016), began compiling and analyzing photos and videos of the bomb sites, many of them publicly available on various social media websites. "Suspicious" individuals were tracked across multiple photos, with inferences

made regarding their movements and motives—a process similar to the behavioral indicators used by the Transport Security Administration (TSA) that have been shown to be nearly totally ineffective and unreliable (United States, 2013). Users drew diagrams to show who was watching the race and who was not, ascribing suspicion to the latter. They described where people were in multiple frames, adding a temporal dimension to this constructed space. They pointed out when "suspects" were carrying dark bags that may have contained the bomb and when it seemed like they no longer had their bags. And, using diagrams, they showed how a pressure cooker bomb, like the one used in the bombings, could fit in certain bags. Hundreds of photos were collected and analyzed, producing lengthy discussions on /r/findbostonbombers. Anyone on the internet with access to Reddit and various image hosting sites could join this process from afar, contributing to the unsolved investigation by offering clues, usually through imaginative interpretations of photographs. Through this process, several key "suspects" initially rose to the top of the subreddit, included two men of color carrying bags. On the morning of April 18th, these two would find themselves on the front page of the New York Post surrounded by giant block letters that read, "BAG MEN: Feds seek these two pictured at Boston Marathon." The Post referred to online sleuths as "investigators probing" the event and in small type admitted that there was no direct evidence linking the suspects to the crime. Of course the supposed suspects had nothing to do with the bombing—they turned out to be a high school runner and his coach—and one cannot help but speculate about how much racist imaginaries of terrorism contributed to this outcome. Indeed, an image from 4chan containing another "suspect" lists four criteria for the suspicion: "1: ALONE, 2: BROWN, 3: Black backpack, 4: Not watching."[2] After the image was posted to Reddit, a number of redditors began commenting on the racist assumptions that appeared to underlie the analysis of images. As one redditor pleaded, "please

stop picking brown people out in a crowd and speculating they're the bomber so CNN doesn't pick it up and false report." While the 4chan example is the only explicitly racist criteria for ascribing suspicion that I could find in my research, and /r/findbostonbombers explicitly forbade racism in its rules, crowdsourced surveillance in its call to report suspicious activity can easily lead to expressions of underlying prejudices (Trottier, 2014: 622).

Many redditors foresaw the potential damage of the effort, some even before the New York Post's reckless cover debacle. One redditor put it succinctly with the observation that, "This entire sub is a witch hunt. Innocent people *are* getting hurt." Another connected the investigation with the much-maligned surveillance state, observing, "I find it ironic that we bemoan the rise of the surveillance state, but then when something like this happens, everyone's more than happy to post pictures all over the internet, drawing big red circles around anyone carrying a backpack. Let's just round up everyone in Boston who was carrying a backpack that day and waterboard them until they tell us about all the pressure cookers." Some countered these criticisms by praising the efficacy of the subreddit and its potential for aiding official investigations through the added labor. As one redditor surmised, "If you were a fly on the wall in any police investigation or news room this subreddit is similar to what goes on," a sentiment that was countered elsewhere by pointing out that redditors were not trained to do investigative work and official investigators kept their persons of interest private, thereby avoiding the potential harm that can come from making innocent people's identities public. Many defenders of the subreddit blamed the media for making certain posts go viral, thereby ignoring the subreddit's rules, which included: "Remember, we are only a subreddit. We must remember where helping ends and the job of professionals begins" and "Do not make any images viral. Limit reposting images outside of this sub," neither of which was closely observed. Some participants pointed out that Reddit *was* the media, and public at that, so one must recognize the complicity of redditors with the spread of

false accusations. In response to a post claiming that "Until the media got involved, none of the images were going anywhere but to the FBI," a redditor wrote, "It's the Internet. The naive notion that this would not spread past 'sending pictures to the FBI' is stunningly ignorant." And rumors emanating from the subreddit did spread, with the New York Post and others echoing speculative discussions happening in the subreddit.

The FBI also recognized the potential harm that these online investigations might produce, especially in light of media outlets using them as source material for speculation, releasing a statement one day before the Post's blunder, which read, "Over the past day and a half, there have been a number of press reports based on information from unofficial sources that has been inaccurate. Since these stories often have unintended consequences, we ask the media, particularly at this early stage of the investigation, to exercise caution and attempt to verify information through *appropriate official channels* before reporting" (FBI, 2013c, emphasis mine). Some sources even speculate that the release of the images of the Tsarnaev brothers—the FBI's real suspects in the case–on April 18th was done prematurely in an attempt to preempt the circulation of rumors emanating from online sources (Montgomery et al., 2013). As it turns out, none of the images analyzed in the original crowdsourced investigations even contained the Tsarnaev brothers. And the misidentifications would continue, with rumors that one of the FBI's suspects was Sunil Tripathi, a missing Brown University student whose dead body would be pulled from the Providence River a few days later, the apparent victim of suicide. The moderator of /r/findbostonbombers would later apologize to the mourning Tripathi family, adding an accusatory note that read, "This event shows exactly why the no personal information until confirmation rule is in place." Other rumors would abound as misleading information culled from Reddit discussions would leak out and become circulated on other networks like Twitter and

Facebook. As one redditor noted in a comment thread reflecting on the event two years later, "It's probably important to note that there were only a small handful of redditors who drew conclusions too quickly. Everything was rejected eventually by the community as a whole (including the sub itself). However, given the organic nature of Reddit, once the individual accusations were out there, no matter how much correction was attempted, it was already too late."

The Entanglement of Algorithms and Cultures

"None of us are as dumb as all of us." —redditor

This appeal to the "organic nature" of the platform implies that it exists as a neutral conduit for communication. This discursive framing is a common strategy of online content providers as a way to elide their role in curating content (Gillespie, 2010). But how content is shaped, sorted, and accessed through the work of algorithms and moderating practices contributes to the political outcomes of that system. In many reflexive discussions about Reddit, the culture of the user base is addressed, including moderating practices, but its algorithmic affordances are ignored. But, as others have argued in regards to Facebook (Bucher, 2012), YouTube (Gillespie, 2010), Reddit (Massanari, 2015b), and social media more broadly (Beer, 2009), how algorithms intersect with practice has important implications for the actions that emerge.

The Reddit platform relies heavily on voting algorithms to sort posts and comments. Each user is allowed one vote per post, either an upvote or downvote, which adds or subtracts a point to the post's total score. Posts with higher scores rise to the top of the subreddits, making them visible to more users, and subsequently decay over time to make room for newer posts. This score also contribute to a user's karma score—an overall measure of the popularity of a user's

contributions to the site. Each post has a comment section—a central feature of the website and the main repository for user-generated content—which relies on a similar system of upvoting and downvoting, without the dependence on the time variable.[3] And while users are able to select from a number of algorithms that sort posts and comments in different ways, chronologically for example, the default option for posts and comments has them sorted by popularity. This causes popular comments and posts to rise to the top.

As Adrienne Massanari (2015a) has argued in her ethnography of Reddit, this can have a "herding" effect, creating what redditors refer to as the "hivemind," as users are influenced by the voting behaviors of the group. She shows how the popularity voting system of Reddit can create a space where a few dominant voices or positions come to dominate most comment threads (2015a: 154). Indeed, it is a common for dissenting, complex, or nuanced posts to begin with, "I know this will be downvoted, but…" indicating the poster expects the post to fall to the bottom of the comment section and out of sight if it receives enough downvotes. For example, on the post apologizing to the Tripathi family, one redditor argued that Reddit will never learn from this debacle, arguing that this kind of thing will happen again and "the few people who will try to argue for restraint will be heavily downvoted."

Similarly, James Surowiecki (2004: 36), who popularized the idea of collective intelligence that prefigured work in crowdsourcing, argues for the need for diversity and independent thinking to avoid the propagation of "groupthink." "One of the quickest ways to make people's judgements systematically biased is to make them dependent on each other for information," he writes, "You can be biased and irrational, but as long as you're independent, you won't make the group any dumber" (2004: 41). In his reflections on Reddit in the wake of the Boston bombing, Surowiecki (2013) argues that the structure of Reddit works against these requirements for successful collective intelligence. He observes that the people were not independent: "you need

people to be thinking for themselves, rather than following the lead of those around them," echoing the findings of Massanari (2015a) and others (Bucher, 2012: 1176; Muchnik et al., 2013). So instead of facilitating a wise crowd that does not "act like a crowd at all," (Howe, 2009: 143) we find something else emerging that looks more like a dangerous or irrational crowd. Indeed, detractors who pointed out the dangers of the subreddit often referred to the redditors that made up r/findbostonbombers as a mob. This framing evokes classical sociological theory that warns of the dangerous potential of crowds as individuals became subordinate to the rule of the suggestible mob and are incited into reckless behaviors (Borch, 2013).

This mob behavior also reflects underlying cultural elements of the site as a whole. As Massanari (2015a: 61) has argued, the imagined community of Reddit, according to many redditors, is often described as being constituted by straight, white, geeky, college-educated young men who embody a culture of 'geek masculinity.' This imaginary of the typical redditor is self-reinforcing as voices that are perceived to deviate from that norm are routinely silenced through downvoting, harassment, and general dismissal (2015a: 156). Massanari (2015b) argues that these 'toxic technocultures' have been able to emerge on Reddit for a number of reasons, including technological affordances of the system, limitations in governance structure, and lack of policies to limit harassment—all factors that are ignored in discursive framings that make appeals to the so-called neutrality of the platform. It is important to note that not all of Reddit embodies this culture—politically progressive, feminist, and anti-racist subreddits exist alongside reactionary, racist, and misogynist ones. These latter voices, however, tend to leak out into the site as a whole, becoming part of the conversation on popular subreddits, thus spreading the toxic environment to other parts of the site (Massanari, 2015a: 138). So, while it may be impossible to ascertain the specific subject positions of the loudest voices in /r/findbostonbombers, we can trace some of the dominant forces that shaped the outcome of this case of online surveillance.

It is important, as Kevin Haggerty (2006) has argued, to remember who operates the surveillance apparatus. Haggerty (2006: 33-34) critiques Foucault's "failure to contemplate the specific characteristics of the operatives conducting surveillance" in his work on the panopticon, thus missing how the masculinist gaze shapes the results. Feminist scholars have recently echoed this deficiency in security scholarship, calling for a feminist surveillance studies that, among other things, considers the embodied contexts of those who surveil and those who are surveilled (Dubrofsky and Magnet, 2015). In the case of Boston, complex subject positions, shaped by algorithmic affordances and a dominant imaginary of white male masculinity, influenced the results of the crowdsourced effort. The culture of Reddit and the affordances of the platform—which emerged in tandem, each shaping the other—structured the possibilities for thought and action following the bombing.

Surveillance Imaginaries

These cultures and affordances also ran up against conflicting surveillance imaginaries. While it is already a difficult task to study open source software like Reddit, most commercial and governmental software logics are intentionally hidden, obfuscated, or otherwise black boxed (Gillespie, 2014; Kitchin, 2014; Langlois and Elmer, 2013). This unknown served an important rhetorical function, leading some redditors to speculate that the FBI software and methods did not provide a significant advantage over the methods used by redditors. As one redditor argued: "Do you think the FBI and other law enforcement have some mystical superpowers at divining the truth? The don't, they have to work hard just like the redditors here have been doing to sift through as much as possible and identify what they can. They may very well have advanced forensics software that helps them keep up with subjects across an area over time, but if so.. it hasn't seemed to do them any good --yet." Others countered by speculating that the FBI's

software was far ahead of any that was publicly available and, in addition, the FBI had much better training and access to far more data. The last claim became verifiably true after images culled from private security cams were released by the FBI in an effort to track down the Tsarnaev brothers. Regardless of the real capabilities of the FBI's software or its access to data, the imaginary of its investigative capacities mediates users' participation in crowdsourced surveillance efforts. This ran the gamut from those who claimed their efforts were needed in the face of an inept government to those who called for the shuttering of the entire subreddit, arguing that all information should only be funneled through the appropriate legal channels.

Surveillance imaginaries become further confused by the rhetoric that both sides of the debate deploy. On one hand, we can point to the successes of untrained internet detectives in leveraging social media for surveillance purposes. These cases often lend credence to claims of the state as inept or unwilling to investigate incidents using the surveillant powers of the internet. For example, in 2011 a UC Davis police officer pepper-sprayed a group of seated, protesting students. Following widespread public outcry after the distribution of videos of the event, the officer was "doxed" by the online group Anonymous, which means they identified him from the videos and released his personal information, including his name, cell phone number, and home address. Subsequently, the officer received tens of thousands of threatening emails, text messages, and letters denouncing his actions and was eventually dismissed (Carroll, 2012). Another example, shortly after Boston, leveraged Facebook's release of Graph Search in 2013, which allows detailed searches of masses of Facebook data. It facilitated the search for suspects in a hate crime against a gay couple in Center City Philadelphia. After the Philadelphia police released surveillance footage of the crime, Twitter users were able to find a photo of the suspects from their Facebook check-in data at a nearby restaurant—information the suspects probably did not realize was publicly accessible (Dewey, 2014). The information led to their arrests and further

research exposed homophobic tweets previously posted by one of the suspects. Philadelphia Police Detective Joe Murray praised the effort, with a tweet that read, "This is how Twitter is supposed to work for cops. I will take a couple thousand Twitter detectives over any one real detective any day." Stories like these contributed to many redditors' earnest belief in the efficacy of their efforts. Some explicitly referenced successful crowdsourcing efforts, including the identification of people suspected of rioting following a hockey game in Vancouver (Schneider and Trottier, 2012). "In Canadian riots the other year they crowdsourced the investigation to discover the identities of many criminals. This process is not unprecedented," observed one redditor in /r/findbostonbombers. These imaginaries, it should be noted, did not go uncontested. For example, in reply to the previous quote, another redditor pointed out an important discrepancy between the two cases, observing, "There is a clear and obvious difference, which is that they crowdsourced the images of *people who were already known to have committed crimes*. They didn't say 'here's some pictures from an hour before the riots, please wildly speculate about which people here might have rioted and send us your guesses.'"

On the other hand, lofty rhetoric by law enforcement paints a picture of a surveillance apparatus far more advanced than anything available to the public. Consider, for example, the following statement by Tim Murphy, former Deputy Director of the FBI, two days after the bombing: "It's an overwhelming task. All that information will be in a repository and they will be able to search across links. And the system itself will make links. Whether it's a person, place, or thing, there are searches that are done to see if this name or this location or this information has come up in other cases. And you can do a google-like search across this information and that will connect the dots for us" (CBS, 2013). In Boston, however, the software did not connect the dots—face recognition software failed to identify the Tsarnaev brothers once they emerged as suspects (Klontz and Jain, 2013). The FBI relied on information provided by a family member to

eventually ascertain their identities. Despite the failure of this software, it is noteworthy that the FBI has access to private databases that hold the possibility of identifying people in a crowd from surveillance footage.[4] This leads some commentators, both within the computer science world (Klontz and Jain, 2013) and the political world to speculate that better software could have connected the dots. These voices become part of popular discourse, lending credence to claims that this event could have even been prevented through the use of the better technology. For example, in a 2015 Republican debate, Carly Fiorina asks why we missed the Tsarnaev brothers. "It wasn't because we had stopped collected metadata it was because, I think, as someone who comes from the technology world, we were using the wrong algorithms," she argues (Team Fix, 2015). If the data had been read in the correct way, or so the argument goes, preemptive logic (Anderson, 2010: 789-90) could have recognized the potential threat that the Tsarnaev brothers posed and intervened before the attacks. If this sounds outlandish, consider nearly identical claims made by security groups in the wake of the 9/11 attacks (Amoore and De Goede, 2005: 160; Grusin, 2010: 123).

These surveillance imaginaries contribute to shaping relationships to and understandings of surveillance. Certainly, Reddit's successes and failures in crowdsourcing surveillance shape how it relates to future efforts. The refrain "We did it, Reddit!" often pops up in comment threads—a sarcastic allusion to the failures of /r/findbostonbombers, which serves as a cautionary reminder to temper claims and accusations. But belief in the power of collective intelligence based on a combination of actual successes and ideological imaginaries also contributes to continued participation in crowdsourced surveillance projects. Alternatively, overblown claims by technologists can lead to magical thinking about computers by the public, thus eliding the actual affordances, possibilities, and limitations of actually existing technologies. These claims can also lead to imaginaries that posit data as puzzles that can be solved under the correct conditions. In a

technologically-savvy community like Reddit, these imaginaries can be a strong motivator for self-described geeks to attempt to solve the puzzle. Of course, how all of this plays out, as described before, is always contextual, hence the need for specific studies. And social media is quite good at revealing the possibilities afforded by this new, vast surveillance apparatus of networked computers, in addition to being of growing interest to law enforcement. For example, in its "2015 Social Media Survey," The International Association of Police Chiefs (2015) found that 85.5% of responding agencies reported that social media had helped them solve crimes. As law enforcement increasingly turns to social media for surveillance purposes, it is important to study its latent possibilities.

Constructing Worlds

"If seeing is believing, it is also techno-culturally mediated" (Gregory, 2011: 203).

Internet users were able to realize additional surveillant possibilities of social media by tapping into the numerous streams of information to construct an understanding of the event as it evolved. Application Programming Interfaces, or APIs, allow users to access data from corporate websites like Facebook and Twitter. Users can tap into these data streams and search social media posts by geolocation, time, and keywords and develop novel connections between distributed data sources. Twitter feeds, image uploads, video streams, live news feeds, police scanners broadcast online, user-generated maps, radio feeds, and forum discussion were all collated as in an attempt to make sense of it all. The fragmentary and often contradictory information that emerged resisted the traditional grand narratives of news reporting, instead inviting internet users to experience and construct the event in the moment, mediated through novel surveillance methods. The availability of certain types of data—EXIF metadata attached to images, which often contain timestamps and geolocations, for example—enabled particular

types of reconstructions. Some redditors argued for the need to move away from the popular image sharing site imgur.com because its algorithms stripped the valuable metadata.

As the manhunt for the Tsarnaev brothers intensified, user-generated and collated content produced multiple ways to experience the event in the moment. Redditors created long news feeds in posts, updated every few minutes with bits of information describing breaking news. Others engaged in lengthy discussions about these feeds, adding commentary, analyses, and more information. Some users began plotting events on google maps, adding a spatial dimension to the story. These users, distributed across the globe, all became part of the event in ways that fed back into the material spaces of Boston in complex ways. The real potential material consequences of this participation was echoed in a plea from the Boston Police Department delivered via Twitter, which read, "WARNING: Do Not Compromise Officer Safety by Broadcasting Tactical Positions of Homes Being Searched." The NY Post's blunder and the cautionary release from the FBI are also both evidence of this feedback. As Reuben Rose-Redwood has argued, "representation has the capacity to reshape the world in its own image" as data and its sorting by software comes to partition the spaces of the world (Rose-Redwood, 2012: 299). Crowdsourced surveillance is not only a matter of watching the world, but also of constructing worlds.

One way worlds are constructed is through algorithmic and surveillance practices that make certain things visible in the world. Social media algorithms distribute visibility in particular ways, making some things "visible, and thus knowable in a specific way" (Bucher, 2012: 1171). This 'distribution of the sensible' delimits not only what is visible and audible in the world, but also "what can be said, thought, made, or done" (Rancière, 2006: 85). As users interact with algorithms, particular elements becoming visible to the community of users (Langlois and Elmer, 2013: 14). In Reddit's search for the Boston bombers, the 'distribution of the sensible' was shaped

by the platform's voting algorithms and moderating practices alongside discursive efforts to ascribe meaning to visual evidence.

Some of these efforts attempted to use visual evidence to produce a complex understanding of the space of the event. One post titled "These are the exact locations of the bombs," posted in /r/findbostonbombers, offered a spatial reconstruction of the bomb sites. It was anomalous within the subreddit in its total lack of discussion about possible suspects. Initiated by a redditor who claimed to work in video analytics and computer vision processing, the post discussed a number of videos and images that contained the two blasts. Overlaid were lines and arrows that merged the various perspectives in an attempt to reveal the physical location of the bombs. These locations could then be marked on photos depicting the aftermath of the blasts. In the comments section, the diagrams and accompanying commentary received a lot of praise for building such a compelling reconstruction of the physical space of the event. Many redditors, it seemed, were tired of the speculative witch hunt that were transpiring across the site, but also thirsted for additional information to reconstruct and understand what had happened. The sheer number of user-generated photos and videos enabled a convincing account of the space of the bombings. Users were able to geolocate various objects and people as they moved in and out of observers' camera frames. This process helped participants resolve the diverse temporal and perspectival views to create a complex understanding of the space of the event.

Algorithmic methods have the potential to automate this process of constructing understandings of space by assembling geolocated data collected through social media APIs. For example, Rashomon,[5] a project of UC Berkeley's CITRIS Data and Democracy Initiative, develops automated methods for users to sync up multiple video feeds captured simultaneously. The goal of the project is "to allow the public to gain a richer understanding of contested events from user-generated video and photo than is currently available online." An implicit claim of the

project is that more data will provide clearer, more incontrovertible narratives of events, especially in response to oppressive tactics of the state.

Other efforts in Boston used visual evidence in an attempt to understand human actions in the moment—categorizing behaviors perceived to be abnormal as suspicious. They hint at visual epistemologies at work as we construct an understanding of the world through images filtered through digital media and visualization technologies. It is interesting to note that some descriptions of official investigations into the Boston bombing closely resemble descriptions of the methods deployed by users on Reddit and 4chan (Montgomery et al., 2013). They also resemble methods used by drone pilots as war is fought from afar. As Derek Gregory writes: "The hierarchies of the network are flat and fluid, its spaces complex and compound, and the missions are executed onscreen through video feeds and chat rooms (displays show as many as 30 different chats at a time) that bring a series of personnel with different skills in different locations into the same zone" (Gregory, 2011: 195). And so just as runners become terrorists in Boston, in drone warfare "objects become rifles, praying a Taliban signifier, civilians 'military-aged males', and children 'adolescents' (Gregory, 2011: 203). If these imaginative constructions consistently reproduce racial profiling, stereotypes, and oppressions, there are no purely technical fixes. So while additional data and images or better algorithms might change how we understand a story, they are always mediated through and built upon particular epistemological frameworks. And all of these factors coalesce to produce subjects as their effect—shifting and unstable subjects mediated through the complex assemblages of crowdsourced surveillance.

An analysis of data points also contributed to the production of subjects in Boston. This process became important as the Tsarnaev brothers emerged as suspects and attempts were made to construct and understand who they were. Internet sleuths uncovered what is thought to be Tamerlan's Amazon wish list, which includes books on how to forge IDs, Chechen history, and

Mafia stories. They also found Dzhokhar's twitter account. One tweet posted in the hours following the bombing reads, "Ain't no love in the heart of the city, stay safe people." Another reads "I'm a stress free kind of guy," which was posted the day before his reemergence that would eventually lead to his arrest. The data trails of the suspects were central to their rendering as subjects in the moment (and it is interesting to read how things like tweets were grossly misconstrued by prosecutors in Dzhokhars trial—many references to jokes and pop culture were lost on lawyers). The subjects in this case are reconstructed after the fact of the event through an imaginative mapping of their past "proclivities and potentialities" (Amoore, 2011: 28). As Deleuze argued, the societies of control, in which we are now immersed, produce the "dividual"—a fractionated subject made up of disaggregated data points (Deleuze, 1992: 5; Amoore, 2013: 9, 92). But how this subject is constructed, as we can see from this case, is highly contingent and unstable.

Conclusion

The strange thrill of remote, distributed surveillance as people search twitter, video feeds, and forum threads to understand an unfolding event has become increasingly commonplace in today's highly-connected world. As one redditor observes in the r/findbostonbombers subreddit, "It's amazing, powerful and intoxicating to get this fresh information on this tragedy." In these moments, complex mechanisms of surveillance, control, and affect coalesce around these software-mediated visualization technologies with unpredictable results.

While redditors were unable to find the Boston Bombers, their actions were a powerful testament to the ability of a group of internet users to track individuals across space and time and attempt to ascertain their identities. Also, it is important to remember that the examples described above all rely on publicly-accessible data, only hinting at the much vaster surveillance

capabilities of groups like the NSA. Subsequent images released by the FBI in the Boston bombing case culled from private video feeds show the Tsarnaev brothers shopping at Target and buying milk at Whole Foods. The ability to access, collate, and sort these private data streams, which follow us through our everyday lives, provides immense possibilities for tracking of individuals. Crowdsourced surveillance examples, even with their limitations of data access, provide us with important first-hand knowledge of the possibilities to not only track subjects, but make inferences about their thoughts, actions, and motives.

Scholars have begun paying attention to the ways that software both enables and mediates remote, distributed participation in events. From mutual monitoring using search engines (Andrejevic, 2004) to remote surveillance of border crossings (Tewksbury, 2012), and disaster relief from afar (Zook et al., 2010), we are witnessing an emerging body of work that explores the implications of crowdsourced surveillance. These cases rely on and produce relations of power that enroll computer users, police, governments, juridical apparatuses, social media corporations, programmers, bystanders, and others in complex ways. They resist simple top-down understandings of surveillance, instead relying on the participation of many people, whether intentionally or not. By capturing our own complicity within these systems, we can begin mapping the complex networks that make surveillance possible and avoid deferring all analyses onto abstract, shadowy "others" like the NSA, whose internal workings are largely inaccessible except through data breaches. They allow us to begin seeing how our own everyday actions and participation in computer networks produces particular power relations and possibilities for surveillance. These mechanisms of surveillance that are produced vary contextually and are constantly shifting with changing technologies and social relations.

The surveillance state that becomes visible does not preexist practice, but is produced through complex processes and the enframing of particular practices (Mitchell, 1999: 89-90),

including some of the ones described above. The state selectively validates some efforts of crowdsourced surveillance, while disavowing others, sometimes because they are ineffective, other times because they challenge state power. As Foucault (2007: 239) once argued, understanding the state as a phenomena "would involve showing the bundle of processes and the network of relations" that constitute it. Take, for example, what Shelton, Zook, and Wiig (2014: 3) call the 'actually existing smart city,' which is "assembled piecemeal, integrated awkwardly into existing configurations of urban governance and the built environment." By understanding how particular practices become enframed within these imaginaries, we can begin to think critically about how they structure thought and action. Specifically, we can begin to understand how particular types of data and algorithms might constitute particular power relations and how these become integrated with existing configurations of governance. So, for instance, when cities propose smart or crowdsourced policing, existing examples might help guide our critical responses to them. In response to automated methods that purportedly circumvent racial profiling practices (Adey, 2009: 284-285), for example, we should not be surprised to find that racial stereotypes can become encoded into surveillance software (Salter, 2006: 182; Barocas and Selbst, 2015). If this software is what structures relationships between people and the police, critical interventions become important.

We might also imagine a myriad of ways that social movements can use this knowledge to intervene in dominant surveillance apparatuses. First, through world-building projects that attempt to mitigate community problems by deploying the tools of surveillance. Consider, for example, crowdsourced sensing systems dedicated to environmental justice (Monahan and Mokos, 2013; Jeremijenko, 2005). Second, by using surveillance to fight against social injustices, as activists "reveal the violence that is more usually buried and concealed beneath the surface" (Amoore, 2013: 126). We have witnessed this recently as the Black Lives Matter movement has

brought police violence into the public spotlight, shifting popular discourse about race and policing. This is also the goal of many campaigns undertaken by the hacker group Anonymous in reaction to perceived injustices around the world (Coleman, 2014). And finally, by providing "bad data" to confuse powerful surveillance apparatuses, either through obfuscation (Brunton and Nissenbaum, 2015) or the use of encryption and proxy servers that render one's data less traceable, for example. With surveillance intervening in the material world in complex ways, we can expect a profusion of new ways to live with and counteract these ubiquitous forces.

References

Adey P (2009) Facing airport security: affect, biopolitics, and the preemptive securitisation of the mobile body. *Environment and Planning D: Society and Space* 27 (2):274–295.

Alexa (2016) *reddit.com Site Overview.* Available at: http://www.alexa.com/siteinfo/reddit.com (accessed 9 February 2016)

Amoore L (2011) Data Derivatives: On the Emergence of a Security Risk Calculus for Our Times. *Theory, Culture & Society* 28, no. 6: 24–43.

Amoore L (2013) *The politics of possibility: risk and security beyond probability*. Durham: Duke University Press.

Amoore L and De Goede M (2005) Governance, risk and dataveillance in the war on terror. *Crime, Law and Social Change* 43 (2-3):149–173.

Anderson B (2010) Preemption, precaution, preparedness: Anticipatory action and future geographies. *Progress in Human Geography* 34 (6):777–798.

Andrejevic M (2004) The Work of Watching One Another: Lateral Surveillance, Risk, and Governance. *Surveillance and Society* 2 (4).

Bhardwaj A (2014) Harnessing the Crowd for Neurology Research. *Science Translational Medicine* 6 (250):250ec141–250ec141.

Barocas S and Selbst A (2015) Big Data's Disparate Impact. *California Law Review* 104. Available at SSRN: http://ssrn.com/abstract=2477899.

Beer D (2009) Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society* 11 (6):985–1002.

Borch C (2013) *The politics of crowds: an alternative history of sociology*. Cambridge: Cambridge Univ. Press.

boyd d (2011) Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. In Papacharissi Z (ed.) *Networked Self: Identity, Community, and Culture on Social Network Sites*. New York: Routledge, 39-58.

Brabham DC (2013) *Crowdsourcing.* Cambridge, Massachusetts ; London, England: The MIT Press.

Brunton F and Nissenbaum HF (2015) *Obfuscation: a user's guide for privacy and protest*. Cambridge, Massachusetts: MIT Press.

Bucher T (2012) Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society* 14 (7):1164–1180.

Carroll R (2012) UC Davis pepper-spray officer fired despite being cleared by internal panel. *The Guardian.* 2 Aug. Available at: http://www.theguardian.com/world/2012/aug/02/uc-davis-pepper-spray-officer

CBS (2013) *CBS This Morning.* 17 April. Available at: https://archive.org/details/KPIX_20130417_140000_CBS_This_Morning#start/3960/end/4020

Coleman EG (2014) *Hacker, hoaxer, whistleblower, spy: the many faces of Anonymous*. London ; New York: Verso.

Deleuze G (1992) Postscript on the Societies of Control. *October* 59:3–7.

Dewey C (2014) How an anonymous Twitter sleuth may have solved a Philadelphia hate crime (and restored our faith in the Internet). *The Washington Post* 17 Sept. Available at: http://www.washingtonpost.com/news/the-intersect/wp/2014/09/17/how-an-anonymous-twitter-sleuth-may-have-solved-a-philadelphia-hate-crime-and-restored-our-faith-in-the-internet.

Dubrofsky R and Magnet S (eds) (2015) *Feminist Surveillance Studies*. Durham: Duke University Press.

FBI (2013a) Statement of Attorney General Eric Holder on the Ongoing Investigation into Explosions in Boston. Available at: http://www.justice.gov/opa/pr/statement-attorney-general-eric-holder-ongoing-investigation-explosions-boston.

FBI (2013b) Remarks of Special Agent in Charge Richard DesLauriers at Press Conference on Bombing Investigation. Available at: https://www.fbi.gov/boston/press-releases/2013/remarks-of-special-agent-in-charge-richard-deslauriers-at-press-conference-on-bombing-investigation.

FBI (2013c) No Arrest Made in Bombing Investigation. http://www.fbi.gov/boston/press-releases/2013/no-arrest-made-in-bombing-investigation.

Foucault M (2007) *Security, territory, population: lectures at the Collège de France, 1977-78*.

Gillespie T (2010) The politics of "platforms." *New Media & Society* 12 (3):347–364.

Gillespie T (2014) The Relevance of Algorithms. In: Gillespie T., Boczkowski PJ and Foot KA (eds) *Media technologies: essays on communication, materiality, and society*. Cambridge, Mass: MIT Press.

Goffey A (2008) Algorithm. In: Fuller M (ed.) *Software studies: a lexicon*. Cambridge, Mass: MIT Press.

Gregory D (2011) From a View to a Kill: Drones and Late Modern War. *Theory, Culture & Society* 28 (7-8):188–215.

Grusin RA (2010) *Premediation: affect and mediality after 9/11*. New York: Palgrave Macmillan.

Haggerty K (2006) Tear down the walls: On demolishing the panopticon. In: Lyon D (ed.) *Theorizing Surveillance: The Panopticon and Beyond*. Cullompton: Willan, 23–45.

Howe J (2006) The Rise of Crowdsourcing. *Wired Magazine*. 1 Jun. Available at: http://www.wired.com/2006/06/crowds/.

Howe J (2009) *Crowdsourcing: why the power of the crowd is driving the future of business*. New York: Three Rivers Press.

International Association of Chiefs of Police (2015) 2015 Social Media Survey Results. Available at: http://www.iacpsocialmedia.org/Resources/Publications.aspx

Jeremijenko N (2005) Feral Robotic Dogs. Available at: https://www.nyu.edu/projects/xdesign/feralrobots/

Kitchin R (2014) Thinking Critically About and Researching Algorithms. The Programmable City Working Paper 5. *SSRN*. Available at: http://www.ssrn.com/abstract=2515786.

Klontz JC and Jain AK (2013) A Case Study of Automated Face Recognition: The Boston Marathon Bombings Suspects. *Computer* 46 (11):91–94.

Langlois G and Elmer G (2013) The research politics of social media platforms. *Culture Machine* 14:1–17.

Lyon D (2006) Tear down the walls: on demolishing the panopticon. In Lyon D (ed.) *Theorizing surveillance: the panopticon and beyond*. Cullompton, Devon: Willan Publishing.

Massanari A (2015a) *Participatory culture, community, and play: learning from reddit*. New York: Peter Lang.

Massanari A (2015b) #Gamergate and The Fappening: How Reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society* 14(7): 1164-1180.

Mitchell T (1999) Society, economy, and the state effect. In Steinmetz G. (ed.) *State/Culture: State-Formation after the Cultural Turn*. Ithaca: Cornell University Press, 76-97.

Monahan T and Mokos JT (2013) Crowdsourcing urban surveillance: The development of homeland security markets for environmental sensor networks. *Geoforum* 49:279–288.

Montgomery D, Horowitz S, and Fisher M (2013) Police, citizens and technology factor into Boston bombing probe. *The Washington Post* 20 April. Available at: http://www.washingtonpost.com/world/national-security/inside-the-investigation-of-the-boston-marathon-bombing/2013/04/20/19d8c322-a8ff-11e2-b029-8fb7e977ef71_story.html.

Muchnik L, Aral S, and Taylor SJ (2013) Social Influence Bias: A Randomized Experiment. *Science* 341 (6146):647–651.

Rancière J (2006) *The Politics of Aesthetics: the distribution of the sensible*. London ; New York: Continuum.

Reeves J (2012) If You See Something, Say Something: Lateral Surveillance and the Uses of Responsibility. *Surveillance & Society, 10*(3-4), 235-248.

Rose-Redwood R (2012) With Numbers in Place: Security, Territory, and the Production of Calculable Space. *Annals of the Association of American Geographers* 102 (2):295–319.

Salter MB (2006) The Global Visa Regime and the Political Technologies of the International Self: Borders, Bodies, Biopolitics. *Alternatives: Global, Local, Political* 31 (2):167–189.

Schneider CJ and Trottier, D (2012) The 2011 Vancouver Riot And The Role Of Facebook In Crowd-Sourced Policing. *BC Studies*, (175), 57-72.

Shelton T, Zook M, and Wiig A (2014) The 'actually existing smart city'. *Cambridge Journal of Regions, Economy and Society*.

Surowiecki J (2004) *The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. New York: Doubleday.

Surowiecki J (2013) The Wise Way to Crowdsource a Manhunt. *The New Yorker.* 23 April. Available at: http://www.newyorker.com/news/daily-comment/the-wise-way-to-crowdsource-a-manhunt

Team Fix (2015) 5th Republican debate transcript, annotated: Who said what and what it meant. *The Washington Post.* 15 December. Available at: https://www.washingtonpost.com/news/the-fix/wp/2015/12/15/who-said-what-and-what-it-meant-the-fifth-gop-debate-annotated/

Tewksbury D (2012) Crowdsourcing homeland security: The Texas virtual borderwatch and participatory citizenship. *Surveillance and Society, 10*(3-4), 249-262.

Trottier D (2014) Crowdsourcing CCTV surveillance on the Internet. *Information, Communication & Society* 17 (5):609–626.

United States Department of Homeland Security n.d., *If You See Something, Say Something.* Available at: http://www.dhs.gov/see-something-say-something (Accessed 1 February 2016).

United States Government Accountability Office (2013) *Aviation security: TSA should limit future funding for behavior detection activities*. Washington, D.C.: United States Government Accountability Office. Available at: http://www.gao.gov/assets/660/658923.pdf.

Young S, Pinkerton A, and Dodds K (2014) The word on the street: Rumor, "race" and the anticipation of urban unrest. *Political Geography* 38:57–67.

Zook M, Graham M, Shelton T, and Gorman S (2010) Volunteered Geographic Information and Crowdsourcing Disaster Relief: A Case Study of the Haitian Earthquake. *World Medical & Health Policy* 2 (2):6–32.

1. My analysis of /r/findbostonbombers focuses on seven comment threads containing 1344 user comments visible in the default view. These threads were posted on the site between April 17-19, 2013 and accessed via the Internet Archive's Wayback Machine (https://archive.org/web/) using snapshots taken between April 18-20, 2013. Content was analyzed for reflexive discussions of the investigation, arguments and discussions of methods, and technical descriptions. The comment threads were chosen for their focus on some or all of these themes. Comments and discussions were then contextualized using information linked to and discussed in the threads, media and government reports, and summative articles and comment threads published after the fact.

All comments were posted and available publicly online at the time of their writing, attributed to users' pseudonyms. The subreddit was eventually set to private, as described in this text, but archives are still publicly available on the Wayback Machine. All quotes drawn from the subreddit are attributed to "redditor" in this article.

2. http://imgur.com/a/sUrnA

3. The default sorting algorithm for comments is called "Best" and relies on statistical methods to correct for an older algorithm that resulted in early comments rising and staying at the top of subreddits. See http://www.redditblog.com/2009/10/reddits-new-comment-sorting-system.html for more details.

4. The FBI's "Next Generation Identification" website claims its database contains 23 million front-facing photos that can be searched against another photo. See: https://www.fbi.gov/about-us/cjis/fingerprints_biometrics/ngi
5. See https://rashomonproject.org/davis/ for a demo that uses the UC Davis pepper-spraying incident.