ALGORITHMIC HARMS BEYOND FACEBOOK AND GOOGLE: EMERGENT CHALLENGES OF COMPUTATIONAL AGENCY

ZEYNEP TUFEKCI*

INTRODUCTION 203
I. GATEKEEPER ALGORITHMS: THE CASE FOR CONCERN AND CHALLENGES 206
A. Algorithmic Harms: Lack of Visibility, Information Asymmetry, and Hidden Influence 207
B. When "Big Data" Reveals More Than Was Disclosed: Computational Violations of Privacy 209
II. ALGORITHMIC HARMS THROUGH COMPUTATIONAL AGENCY: TWO CASE STUDIES 213
A. Algorithmic Harms: Social Movements and the Civic Sphere 213
B. Elections And Algorithmic Harms 215
CONCLUSION 216

INTRODUCTION

In June of 2014, experiments performed on Facebook users without their knowledge—a previously little-known topic—became big news, garnering enormous coverage nationally and internationally. The furor was sparked by a study published in the prestigious *Proceedings of the National Academy of the Sciences* (*PNAS*). Facebook's researchers had, according to their abstract, confirmed that they had shown "experimental evidence for massive-scale contagion via social networks" by reducing the number of positive or negative posts shown to Facebook users by experimentally manipulating users' algorithmically curated "News Feed."[1] Facebook employee Adam D.I. Kramer was the first author; the other authors were academics—

*Assistant Professor, University of North Carolina—School of Information & Library Sciences.

1. Adam D.I. Kramer, Jamie E. Guillory & Jeffery T. Hancock, *Experimental Evidence of Massive Scale Emotional Contagion Through Social Networks*, 111 PROC. NAT'L ACAD. SCI 24, 8788–90 (2014), http://www.pnas.org/content/111/24/8788.full.pdf.+html.

Jeffrey T. Hancock, a professor at Cornell, and Jamie Guillory, who was Hancock's graduate student at the time the research was undertaken.

Facebook's algorithmically curated News Feed decides which of the "status updates" a Facebook user sees from his or her friends. The researchers positively showed that news and updates on Facebook influence the tenor of the viewing Facebook-user's subsequent posts—and that Facebook itself was able to tweak and control this influence by tweaking the algorithm.[2] While this was not a surprise to anyone who studied Facebook or other online social platforms, the confirmation of the effect, and the mode of the confirmation—through experimental manipulation of the algorithm—sparked a large conversation.

After the media picked up news of the study, a massive outpouring of alarm occurred, likely fueled in part by the very networking abilities platforms like Facebook enable.[3] After controversy erupted, *PNAS* appended an "Editorial Expression of Concern and Correction" to the paper, which expressed concern about the methods and ethics of the study, but did not retract the paper.[4] The emotional contagion article became the most discussed scholarly article on the Internet—ranked #1 of 32,992 scientific articles by Altmetric—as a result of the furor.[5] Hundreds of people upset with the study contacted the lead researcher, who issued an apology about the "anxiety [the paper] caused."[6]

Even Facebook's Chief Operating Officer, Sheryl Sandberg, was confronted with questions on this topic, including whether Facebook might algorithmically throw elections[7]—a possibility which, to the alarm of activists and some academics, was revealed in an earlier separate study.[8] As Facebook's broad effects were again highlighted,

2. *Id.*

3. Reed Albergotti & Elizabeth Dwoskin, *Facebook Study Sparks Soul-Searching and Ethical Questions*, WALL ST. J. (June 30, 2014, 9:01 PM), http://www.wsj.com/articles/facebook-study-sparks-ethical-questions-1404172292.

4. Reed Albergotti & Elizabeth Dwoskin, *Facebook Study Sparks Soul-Searching and Ethical Questions*, WALL ST. J. (June 30, 2014, 9:01 PM), http://www.wsj.com/articles/facebook-study-sparks-ethical-questions-1404172292.

5. Cat Chimes, *Altmetric Top 100 2014*, ALTMETRIC (Dec. 9, 2014), http://altmetric.com/blog/altmetric-top-100-2014/.

6. Facebook Status of Adam D.I. Kramer, FACEBOOK (June 29, 2014), https://www.facebook.com/akramer/posts/10152987150867796.

7. Interview by Schweta Rajpal Kohli with Sheryl Sandberg, COO, Facebook, in New Delhi, India (Updated Jul. 3, 2014), *available at* http://www.ndtv.com/world-news/will-not-try-to-control-emotions-or-elections-facebook-coo-sheryl-sandberg-tells-ndtv-full-transcrip-583183.

8. Micah L. Sifry, *Facebook Wants You to Vote on Tuesday. Here's How it Messed With Your Feed in 2012*, MOTHER JONES (Oct. 31, 2014, 6:00 AM),

the earlier research on elections recaptured the imagination of some critics. Facebook, in its official company statement on the emotional contagion experiment, stated that:

> [t]his research was conducted for a single week in 2012 and none of the data used was associated with a specific person's Facebook account. We carefully consider what research we do and have a strong internal review process. There is no unnecessary collection of people's data in connection with these research initiatives and all data is stored securely.[9]

While the Facebook emotional contagion study was the focus of a media whirlwind, with questions about the ethical and legal status of the experiment garnering much attention, the experiment revealed a greater underlying issue: the ethics of algorithmic and experimental manipulation in online platforms, especially those like Facebook which play an oversized role in directing content, including news and civic information.[10]

On many platforms, including Facebook, algorithmic manipulations are performed routinely; they range from purposes as mundane as deciding the color of a button, to decisions as significant as which news article is shown to the public. These decisions are increasingly central to social, political, and civic processes; consequently, the algorithmic harms that may arise from such computational form a brand new category, including a spectrum of issues from libel to violence. What are the rules around automated code that posts to Twitter accounts (Twitter "bots")? Can a Twitter bot make a death threat?[11] Can it buy illegal drugs or break the law? These questions have already moved beyond hypotheticals, and government authorities are attempting to come to terms with them.[12]

---

http://www.motherjones.com/politics/2014/10/can-voting-facebook-button-improve-voter-turnout; Zeynep Tufekci, *Engineering the Public: Big Data, Surveillance and Computational Politics*, 19 FIRST MONDAY 7 (July 7, 2014), http://journals.uic.edu/ojs/index.php/fm/article/view/4901/4097; Jonathan Zittrain, *Engineering an Election*, 127 HARV. L. REV. F. 335 (Jun. 20, 2014); Jonathan Zittrain, *Facebook Could Decide an Election Without Anyone Ever Finding Out*, NEW REPUBLIC (Jun. 1, 2014), http://www.newrepublic.com/article/117878/information-fiduciary-solution-facebook-digital-gerrymandering.

9. James Vincent, *Facebook Defends Secretly Manipulating Users: Experiments 'Improved Service'*, INDEPENDENT (June 30, 2014), http://www.independent.co.uk/life-style/gadgets-and-tech/facebook-defends-secretly-manipulating-users-emotions-says-experiments-improved-our-service-9572145.html.

10. Tufekci, *supra* note 8.

11. Dan Moren, *Amsterdam Police Respond to Death Threat Created by a Twitter Bot*, POPULAR SCIENCE (Feb. 11, 2015), http://www.popsci.com/amsterdam-police-respond-bot-created-twitter-threat.

12. Tim Cushing, *Twitter Bot "Issues" Death Threat, Police Investigate*, TECHDIRT (Feb. 12, 2015), https://www.techdirt.com/articles/20150211/14002829991/twitter-bot-

In this article, I outline some of the challenges of grappling with algorithmic harms, especially in cases where algorithms act as de facto gatekeepers of consequence.

I. GATEKEEPER ALGORITHMS: THE CASE FOR CONCERN AND CHALLENGES

Algorithms, or computational processes that are used to make decisions, are often deployed as gatekeepers; in this function, they are somewhat similar to the role of a newspaper editor, but possess important differences from their offline, non-interactive and non-computational counterparts. Hence, algorithmic gatekeeping raises significant yet novel issues in many realms.[13]

Algorithms are computer programs, a set of instructions for carrying out procedures step-by-step, and range from quite simple to very complex. While the term "algorithm" is imprecisely broad, it has come to be used for the subcategory of computational processes discussed in this and similar articles.

When I use the word "algorithms" in this article, it is in reference to computational processes that are used to make decisions of such complexity that inputs and outputs are neither transparent nor obvious to the casual human observer. This combination raises questions of accountability, process, and transparency. Further, in this article, I focus specifically on algorithms that do not result in simple, "correct" answers—instead, I focus on those that are utilized as subjective decision makers.

Many algorithms are used for processes where there is a "correct" answer, for example alphabetically sorting a database of names, or calculating the average sales per employee. In such cases, given the inputs and outputs, there is a right answer that is generally not in contention, and the exact algorithmic method for this computation is rarely an issue worth close examination, except in certain edge cases.

In subjective decision-making, there is no such "correct" answer with which to anchor and evaluate the algorithm's operations. For example, consider the study that prompted this article, where Facebook changed the outputs of its algorithmic processes to select which status updates people are shown from their Facebook "friends" and acquaintances. There is no single correct—or even obvious—result from this algorithm that would perfectly apply to all of Facebook's billion-plus users. This complex decision requires "thousands of metrics," containing many moving parts, and is changed

issues-death-threat-police-investigate.shtml; Kashmir Hill, *Who do we blame when a robot threatens to kill people?*, FUSION (Feb. 15, 2015), http://fusion.net/story/48656/who-do-we-blame-when-robots-threaten-to-kill-people/.

routinely (almost every week).[12] Such subjective decision-making by algorithms raises a different set of questions.

Theoretically, it may be best to conceptualize algorithms as "actants," a term from actor-network-theory, in that they are computational agents that are not alive, but that act with agency in the world[14] (though even that term is imprecise applying to this novel category of actors). We can refer to this new category as computational agency. The algorithms and computational agency discussed in this paper differ from other, more straightforward procedures through the combination of their labyrinthine components, processes, and opacity, and the subjective nature of the decisions reached.

It is hard to find an exact offline analogy to the computational agency behind Facebook's filtering of social interactions. One example might be to imagine that your service provider has tasked your smartphone—armed with detailed information about you—with keeping you "engaged" in conversations in order to serve you ads. Your phone would also "decide" which of your friends and family members were the most successful in keep you on the line, and which ones caused you to hang up faster. What's more, your phone silenced calls from those it deemed less "engaging," announced calls with longer rings longer for those it deemed more "engaging," and rearranged speakers' sentences and stories on the fly, as its algorithmic processes served as gatekeeper, with varying degrees of success. None of this is visible to you, or in your control.

In this analogy, your phone would algorithmically manipulate who you heard from, which sentences you heard, and in what order you heard them—keeping you on the phone longer, and thus successfully serving you more ads. Now imagine that the phone company had published an article pointing out that letting more phone calls from your cheerier friends and family members made you cheerier, and those from more gloomy ones resulted in gloomier conversations. You would expect a strong public reaction. Computational agency raises such issues, but with less visibility.

A. Algorithmic Harms: Lack of Visibility, Information Asymmetry and Hidden Influence

Algorithmic gatekeeping is the process by which such non-

11. *See, e.g.*, Ravi Somaiya, *How Facebook Is Changing the Way Its Users Consume Journalism*, N.Y. TIMES (Oct. 26, 2014), http://www.nytimes.com/2014/10/27/business/media/how-facebook-is-changing-the-way-its-users-consume-journalism.html.

14. BRUNO LATOUR, REASSEMBLING THE SOCIAL: AN INTRODUCTION TO ACTOR-NETWORK-THEORY (Oxford 2007).

transparent algorithmic computational-tools dynamically filter, highlight, suppress, or otherwise play an editorial role—fully or partially—in determining: information flows through online platforms and similar media; human-resources processes (such as hiring and firing); flag potential terrorists; and more.[15] The spread of algorithmic gatekeeping is the result of increasing digitization of everyday life, ranging from the personal to the political. This data generated by this digitization, in turn, powers the algorithms that play increasingly crucial roles. Overall, the trend towards more massive datasets (sometimes dubbed "big data"), digital connectivity, and algorithmic gatekeeping are inexorably linked.[16]

Algorithms do not only make decisions in social and interpersonal spheres, but also in political and civic realms. The algorithms' editorial decision-making resembles some traditional editorial processes in many ways, but it differs in many ways too. In newspapers, journalists write stories, which are considered and altered by fact-checkers, editors, and copyeditors. Editors decide whether a story fits or is worthy of publication. If so, the article is published; if not, it is not published. Similarly, Facebook's News Feed and other such algorithmic decision makers "decide" whether a news article shared by one of its users is shown to other users or not.

However, the visibility and transparency differ significantly between these two processes. In Facebook's case, the algorithmic "editing" is dynamic, all but invisible, and individually tailored. Conversely, in print newspapers the result of the editing process is frozen in the newspaper's pages, and is visible to the public—everyone who picks up the same edition of a newspaper sees an identical product, which makes it easier to challenge. Even television—a medium in which journalists, anchors, and those appearing live can speak on the fly—is similar to newspaper in that the results of the gatekeeping process (the broadcast) are visible to anyone who turns to that channel. Again, this facilitates accountability. Further, in print and broadcast there is a clear "chain of command," in that people who made decisions can be held to account later (even if only theory). The argument here is not that newspaper or TV editorial processes are perfect, or desirable going forward, but that algorithms introduce new obstacles in the quest for accountability and transparency

---

15. *See* Nicholas Diakopoulos, *Algorithmic Accountability*, 3 DIGITAL JOURNALISM (forthcoming 2015), *available at* http://www.tandfonline.com/doi/pdf/10.1080/21670811.2014.976411; Tarleton Gillespie, *The Relevance of Algorithms, in* MEDIA TECHNOLOGIES: ESSAYS ON COMMC'N, MATERIALITY, AND SOCIETY 167 (Tarleton Gillespie et al. eds., 2014), *available at* http://6.asset.soup.io/asset/3911/8870_2ed3.pdf.

16. Gillespie, *supra* note 15; Tufekci, *supra* note 10.

in consequential gatekeeping.

With algorithmic gatekeeping, it may never be clear why a specific code acted exactly as it did. Previously, a cable TV channel would have targeted a demographic group, and might make editorial decisions with that group in mind. But the producers would not be able to tweak this broadcast for each person, all the while concealing the tweaks, because the same broadcast would be visible to the channel's entire audience. It is the complex computational process that makes the decision.

It is important that gatekeepers acting with computational agency are able to tweak the content viewers receive on an individualized basis, without being visible. This functionality is often largely unknown to the users of given services. For example, a recent study showed that even at an elite university, sixty-two percent of undergraduates were not aware that Facebook curated users' News Feeds by algorithm—much less the way in which the algorithm works (which is not public information and is subject to near-constant changes).[17] Algorithms are able to act as stealthy, extremely potent gatekeepers; gatekeepers unaccompanied by transparency and visibility.

Hence, while algorithmic gatekeeping performs some traditional gatekeeping functions, it reverses or significantly modifies other key features of traditional gatekeeping with regard to visibility, information asymmetry, and the ability of the public to perceive the results of editorial work.

B. When "Big Data" Reveals More Than Was Disclosed: Computational Violations of Privacy

The privacy, surveillance, and civil rights implications of big data have all recently become the focus of increased scrutiny. Most reports focus on the misuse of data originally disclosed by the user, or on the aggregation of data by entities such as data brokers.[18] These

17. Motahhare Eslami et al., *"I Always Assumed that I Wasn't Really that Close to [Her]": Reasoning About Invisible Algorithms in the News Feed*, PROC. OF THE 33RD ANNUAL SIGCHI CONFERENCE ON HUM. FACTORS IN COMPUTING SYS. (2015), *available at* http://social.cs.uiuc.edu/papers/pdfs/Eslami_Algorithms_CHI15.pdf.

18. *E.g.*, Lois Beckett, *Everything We Know About What Data Brokers Know About You*, PROPUBLICA (June 13, 2014, 1:59 PM), http://www.propublica.org/article/everything-we-know-about-what-data-brokers-know-about-you; FED. TRADE COMM'N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY (2014), *available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf [hereinafter DATA BROKERS REPORT]; AARON RIEKE, DAVID ROBINSON & HARLAN YU, CIVIL RIGHTS, BIG DATA, AND OUR ALGORITHMIC FUTURE (2014),

are certainly important considerations; aggregation of data from large numbers of different sources can create novel privacy threats. However, there is an often overlooked element: developments in computational processes that have grown alongside big data now allow inferences about private information that may never have been disclosed to an online platform. This element is essential to understanding the threats posed, and issues raised, by algorithmic gatekeeping.

For example, using only "Facebook Likes" (a comparative tip of the iceberg in terms of data available on individuals), researchers were able to fairly reliably "model" (computationally and statistically guess to a high degree of accuracy) "latent" traits of 58,000 volunteers.[19] The traits modeled—often with eighty to ninety percent accuracy—included "sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender" among others.[20] It is important to remember here that the algorithm had no other information about these people, or their traits. In other words, it was able to fairly accurately guess if the Facebook user was gay, even if that trait was completely undisclosed on the profile.

This kind of "guessing" via data modeling—where algorithmic processes model data to make a reasonable guess at a trait that is not known or disclosed at all—is in technical literature sometimes called "latent trait inference." Marketers, political strategists, and similar actors have always engaged in such "guessing." Knowing an individual is a NRA member increases the predictive likelihood that she will also vote Republican; knowing an individual is a Planned Parenthood contributor increases the predictive likelihood that she will also vote Democrat.[21] But such categorization has always been imprecise, and beset by a large number of false negatives and false positives.

However, personalized data has grown truly exponentially in just a few years, allowing for great improvements in the computational models used to accurately "guess" such traits. The results turn into more data to feed into models, which in turn can be used to improve the model.[22] The crude approximations of the past always had

http://bigdata.fairness.io/.

19. Michal Kosinski, David Stillwell & Thore Graepel, *Private Traits and Attributes are Predictable from Digital Records of Human Behavior*, 110 PROC. NAT'L ACAD. SCI. 5802 (2013).

20. *Id.*

21. MICAH L. SIFRY, THE BIG DISCONNECT: WHY THE INTERNET HASN'T TRANSFORMED POLITICS (YET) (2014).

22. Lee Gomes, *Facebook AI Director Yann LeCun on His Quest to Unleash Deep Learning and Make Machines Smarter*, IEEE SPECTRUM, (Feb. 18, 2014, 2:45 PM), http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/facebook-ai-

substantial room for error: some Republicans support Planned Parenthood, and some NRA members are Democrats. The current state-of-the-art in latent trait inference, however, goes beyond lumping people into rough categories and uses individualized data—sometimes thousands of data points—to create a deep profile of a person, with increasing levels of accuracy, improving every month. Crucially, such latent trait modeling goes far the beyond basic demographics such as race, gender or social class (which have been available to data brokers for decades) but also includes psychological traits and states of mind. This may now also be coupled with in-depth data about the browsing, informational, civic, and personal interactions of a person, through the merging of online and offline data, or "onboarding."[23] In essence, our machines, armed with our data, can increasingly figure things out about us beyond any previous level, and completely unaccounted for in law, policy or even basic awareness among the general public.

What this kind data analysis can reveal about a person is not only far reaching; it may include information that the person herself did not know. Such methods are not limited to online environments. In a widely publicized example, the *New York Times* reported how a father was upset that Target, the retail giant, was sending his teenage daughter pamphlets of baby products, seemingly encouraging her to become pregnant. When the father complained to the manager of his local store, the irate father received profuse apologies, only to return later to apologize himself: Target's analytics had correctly inferred what he had not known, that his daughter was pregnant.[24]

In another example, researchers were able to identify people with a high likelihood of lapsing into depression before the onset of their clinical symptoms.[25] This type of research is often performed for the best of intentions—the motivation for identifying depression from social media data stems arises from a desire to create early intervention programs, especially in the case of post-partum depression and similar under-diagnosed, life-threatening illnesses. Howev-

director-yann-lecun-on-deep-learning.

23. DATA BROKERS REPORT, *supra* note 18.

24. Charles Duhigg, *How Companies Learn Your Secrets*, N.Y. TIMES (Feb. 16, 2012), http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html.

25. Munmun De Choudhury, Scott Counts & Eric Horvits, *Predicting Postpartum Changes in Emotion and Behavior via Social Media*, MICROSOFT RESEARCH (2013), *available at* http://research.microsoft.com/en-us/um/people/horvitz/predicting_postpartum_changes_chi_2013.pdf; Munmun De Choudhury, Michael Gamon, Scott Counts & Eric Horvitz, *Predicting Depression via Social Media*, *in* PROC. OF THE SEVENTH INT'L AAAI CONFERENCE ON WEBLOGS AND SOCIAL MEDIA 128 (2013), *available at* http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6124/6351.

er, once the knowledge is out of the bag, there is no reason to think it will be used for benign purposes in every instance. Consider an article in *Ad Weekly* spelling out that women are more likely to buy cosmetics when they feel "fat, ugly, lonely" or when they are feeling depressed, colorfully displayed on an informatic chart highlighting the times of the week when women are more likely to feel such emotions.[26] It is not too long a step from identifying women who feels "fat, ugly, lonely" or "depressed" based on the day of the week and time of day, to identifying which ones specifically feel that way in a particular moment, regardless of time of day, and privately and non-transparently targeting them. In addition, the Facebook study showed that Facebook is able to induce mood changes, at least as reflected by postings, in their users through algorithmic manipulation. It is not inconceivable to imagine an advertiser pondering if they can use such fine-tuned, and scientifically shown as effective methods, to induce moods to make people more likely to consume their products. This is something advertisers have always tried to do, but never were able to carry out on an individualized basis; targeting people specifically, and privately, through their vulnerabilities rather than reaching out to broad categories of people all at once.

The main concern is whether this data always gets the right answer: there are a whole host of known biases and problems with big data analyses.[27] The reality remains that the "answer" derived from big data will be used in a variety of cases, raising a set of issues when it is correct, and another set when wrong (especially since we often cannot separate the two cases).

The problem is not that corporations are necessarily undertaking such direct manipulation currently, but that the combination of big data and algorithmic gatekeeping certainly allows this possibility. Given the stakes and rewards, the temptations to implement this should be considered as a real issue.

26. Lucia Moses, *Marketers Should Take Note of When Women Feel Least Attractive*, ADWEEK (Oct. 2, 2013), http://www.adweek.com/news/advertising-branding/marketers-should-take-note-when-women-feel-least-attractive-152753.

27. Zeynep Tufekci, *Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls*, EIGHTH INT'L AAAI CONFERENCE ON WEBLOGS & SOCIAL MEDIA (June 16, 2014), *available at* https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8062.

II. ALGORITHMIC HARMS THROUGH COMPUTATIONAL AGENCY: TWO CASE STUDIES

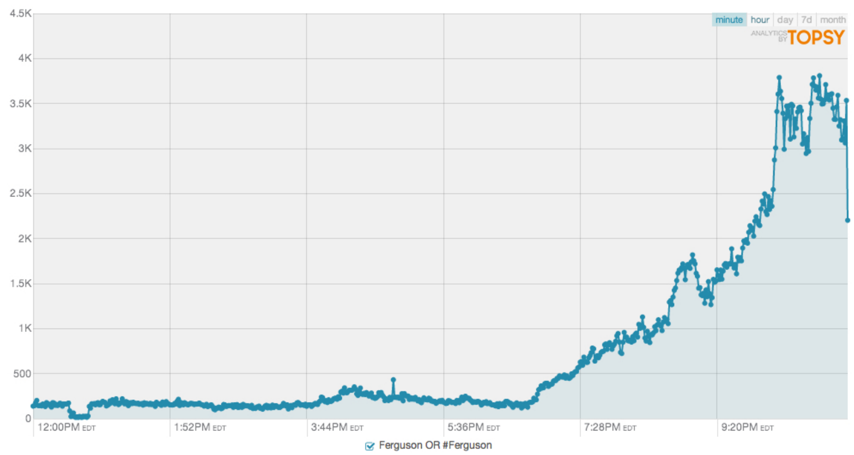
A. Algorithmic Harms: Social Movements and the Civic Sphere

Consider the case of the Ferguson protests, which were triggered after a police officer killed an African-American teenager in Ferguson, Missouri, in August 2014, which in turn, later sparked ongoing nation-wide demonstrations about racial inequalities, the criminal justice system, and police behavior in minority neighborhoods. As the initial protests were percolating, I documented that Facebook's News Feed was algorithmically suppressing news of the protests, then a local event brewing under the radar of national media attention. Acting through computational agency, Facebook's algorithm had "decided" that such stories did not meet its criteria for "relevance"—an opaque, proprietary formula that changes every week, and which can cause huge shifts in news traffic, making or breaking the success and promulgation of particular stories or even affecting whole media outlets. By contrast, Twitter's algorithmically unfiltered feed allowed the emergence of millions of tweets from concerned citizens, which then brought the spotlight of the national media. Algorithmic filtering also by Twitter might have meant that a conversation about police accountability and race relations that has since shaken the country might never have made it out of Ferguson.

The protests started small and local. Michael Brown, a teenager, was shot and killed by a Ferguson police officer, later identified as Darren Wilson. His body had been left in the street for hours, and residents were upset and concerned that an unarmed teenager had been killed. When the local police department in Ferguson aggressively showed up with dogs at the first vigils for this young man, the outrage spread to people who may not have been following the issue on the first day, and started to attract media attention. When reports of tear gas during the nightly protests started pouring in, more national journalists travelled to Ferguson. Some local residents started livestreaming video as well.

On the evening of August 13, the police showed up in armored vehicles with military gear and snipers in position pointing guns at the protesters. That is when I personally started tweeting about it, amazed at the overuse of police force in a suburban area in the United States. Later that evening, as the protests grew tenser, and the police presence escalated even further, two national journalists were arrested while sitting at a McDonald's charging their phones. Their arrest caused even more outrage, and focused the attention of many other journalists, and ordinary citizens on Ferguson.

According to Topsy, a Twitter analytics firm, tweets about Ferguson spiked right around that time.[28]



However, as these events occurred, I found little or no news about Ferguson on Facebook, even though many of my Facebook friends were interested in such events. I started asking on Twitter about other people's experience as they viewed the News Feed. Later, I performed a Twitter search for the keywords "Twitter Facebook Ferguson" and found that hundreds of ordinary people were complaining of a similar information blackout on Facebook, that was instead dominated by the "ice bucket challenge" in which people poured buckets of ice water on themselves in support of a charity, and invited their network to the same.

Traditional media did not go live covering Ferguson until well after the tweets had peaked. At one point, a Ferguson livestream video had about 40,000 viewers, which is about ten percent of the nightly average on CNN at that hour.[29] Two seemingly different editorial regimes, one algorithmic and one edited by humans, had converged to an effective suppression of an issue that would later be pushed onto the national stage. The emergence of Ferguson as an issue on traditional national news media occurred only after sustained protests, as well as Twitter activity, which appears chronologically on a user's feed, had forced the issue on the agenda. Without Twit-

---

28. *See Trending Topics on Twitter on August 13, 2014*, TOPSY, http://topsy.com/s?q=august%2013%2C%202014 (last visited Mar. 29, 2015).

29. Tweet of Trevor Timm, (Aug. 13, 2014 9:22 pm), https://twitter.com/trevortimm/status/499742916315582464 (quoting Mike Brown, *I am Mike Brown Live from Ferguson, MO*, ARGUS STREAMING NEWS (Aug. 10, 2014 1:00 PM), http://new.livestream.com/accounts/9035483/events/3271930).

ter's reverse chronological stream, unmediated by an algorithmic gatekeeper, that allows its users to amplify content as they choose, the news of unrest and protests may never have made it onto the national agenda. Given that so many protests and social movements depend on new media, especially to circumvent censorship and to organize,[30] it is especially important to pay attention to the role algorithmic connectivity plays in civic ecology.

B. Elections and Algorithmic Harms

Algorithmic gatekeeping raises issues far beyond social interaction. Consider the case of elections and voter turnout. In 2010, a massive experiment (performed without being noticed by any of the sixty-one million subjects, none of whom were asked for permission), Facebook demonstrated that it could alter the U.S. electoral turnout by hundreds of thousands of votes, merely by nudging people to vote through slightly different, experimentally manipulated, get-out-the-vote messages.[31] In this experiment, some messages geared toward Facebook users appeared stand-alone while other, more potent ones, were socially embedded, showing a "your friend voted" extra nudge. By later matching the names with voter rolls, thus verifying who had voted and who had not, the researchers found that the social message was significantly more potent than the informational message in nudging people to vote.

Facebook eventually published this research in *Nature*. However, similar research that Facebook undertook in 2012 still has not been published, despite considerable demand including numerous news articles calling upon Facebook to do so.[32] Facebook has stated explicitly that they had tried to keep their 2010 experiment from skewing the election.[33] However, had Facebook not published the re-

---

30. Zeynep Tufekci, *Social Movements and Governments in the Digital Age: Evaluating a Complex Landscape*, 68 J. INT'L AFF. 1 (Dec. 12, 2014), http://jia.sipa.columbia.edu/social-movements-governments-digital-age-evaluating-complex-landscape/; Zeynep Tufekci, *"Not This One:" Social Movements, the Attention Economy, and Microcelebrity Networked Activism*, 57 AM. BEHAV. SCI. 848 (2013), *available at* http://doi.org/10.1177/0002764213479369.

31. Robert M. Bond, Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jamie E. Settle & James H. Fowler, *A 61-million-person Experiment in Social Influence and Political Mobilization*, NATURE (Sep. 12, 2012), http://www.nature.com/nature/journal/v489/n7415/pdf/nature11421.pdf; Alicia M. Cohn, *Study: Facebook Increases Voter Turnout*, HILL (Sept. 12, 2012, 7:27 PM), http://thehill.com/blogs/twitter-room/other-news/249087-study-facebook-increases-voter-turnout; John Markoff, *Social Networks Can Affect on Voter Turnout, Study Says*, N.Y. TIMES (Sept. 12, 2012), http://www.nytimes.com/2012/09/13/us/politics/social-networks-affect-voter-turnout-study-finds.html.

32. Sifry, *supra* note 8.

33. Bond, *supra* note 31. *See also* Tufekci, *supra* note 8; Zeynep Tufekci & Brayden

sult, and had they intended to shape the electorate to favor one candidate over another, the algorithmic gatekeeping enabled through computational agency would have been virtually unnoticeable, since such algorithmic manipulation is neither public, nor visible, nor easily discernible. It is entirely plausible that Facebook can decide close elections, especially in first past-the-post electoral systems where marginal differences make greater impacts. In this particular case, it is a net good that Facebook has published this research, and it should be encouraged to be a responsible actor and continue to do so. However, the fact remains that the general public has little knowledge or control in this important new area.

This threat should be understood within the broader environment in which big data exists, and can be used to infer "latent traits," i.e. private information about people, as explained in the previous section. Facebook could have chosen to tailor these "get out the vote" messages based on inferred personality types and other traits, or otherwise manipulate the message for maximum impact. Since Facebook tweaks its algorithm constantly, a change of this kind might never be noticed since, unlike television, there is no single place where all this can be observed. Also, combining the lessons of the "emotion manipulation" or "massive contagion" study with the voter turnout study, it is clear that Facebook need not use an experimental message to influence turnout. Many people voluntarily post whether they have voted or not, and whether these messages are seen by other people is solely determined by Facebook's News Feed algorithm. Through computational modeling, Facebook can easily suss out the political preferences of its users, even if the users have never indicated it explicitly in their Facebook profiles, or their interactions. From there, it is relatively straightforward to show encouraging, voter friendly status updates to preferred demographic groups, while suppressing those perceived to be less desirable for a particular voting outcome.[34]

CONCLUSION

Algorithms that make decisions open up the same host of ques-

King, *We Can't Trust Uber*, N.Y. TIMES (Dec. 7, 2014), http://www.nytimes.com/2014/12/08/opinion/we-cant-trust-uber.html; Zeynep Tufekci, *Beware the Smart Campaign*, N.Y. TIMES (Nov. 16, 2012), http://www.nytimes.com/2012/11/17/opinion/beware-the-big-data-campaign.html. For more on how elections can be manipulated stealthily, *see* Jonathan Zittrain, *Facebook Could Decide an Election Without Anyone Ever Finding Out*, NEW REPUBLIC (June 1, 2014), http://www.newrepublic.com/article/117878/information-fiduciary-solution-facebook-digital-gerrymandering.

34. Jonathan Zittrain, *Engineering an Election*, 127 HARV. L. REV. F. 335 (2014); Zittrain, *supra* note 33.

tions we have for humans making decisions: transparency, accountability, discrimination, error, and so forth. Computation is increasingly being used to either directly make, or fundamentally assist, in gatekeeping decisions outside of online platforms. Many companies, for example, are turning to algorithms in hiring, firing or other consequential decisions. There are federal laws that apply to many of these decisions, for example, in case of discrimination based on race, gender, disability, family size, or other protected categories in the cases of hiring, loans, or housing. However, algorithmic decision-making creates potential for these types of discrimination to sneak back in, even when not explicitly and directly targeted by algorithms.[35]

For example, rather than race, a hiring algorithm could discriminate based on correlates of race, which would result in a workforce that excluded certain racial backgrounds. This could be done by hiring people based on "commuting distance to work," a factor that companies working on algorithmically calculating the potential success of newly hired employees have already found to be correlated to a low-degree of employee turnover. Such a criterion would not directly target race, but given the residential segregation patterns in many cities around the United States, could easily effectively do so.

The examples provided in this paper are the tip of the iceberg. Computational agency is expanding into more and more spheres. Complex, opaque and proprietary algorithms are increasingly being deployed in many areas of life, often to make decisions that are subjective in nature, and hence with no anchors or correct answers to check with. Lack of external anchors in the form of agreed-upon "right" answers makes their deployment especially fraught. They are armed with our data, and can even divine private information that we have not disclosed. They are interactive, act with agency in the world, and are often answerable only to the major corporations that own them. As the internet of things and connected, "smart" devices become more widespread, the data available to them, and their opportunities to act in the world will only increase. And as more and more corporations deploy them in many processes from healthcare to hiring, their relevance and legal, political and policy importance will also rise. This is a novel challenge, with little to draw from human history as either a metaphor or an analogy. Significant public debate, new ways of thinking, and new approaches will be necessary.

35. Danielle Keats Citron & Frank A. Pasquale, *The Scored Society: Due Process for Automated Predictions*, 18 WASH. L. REV. 1 (2014).