Commentary

Beyond mystery: Putting algorithmic accountability in context

Elizabeth Reddy[1] ⓘ, Baki Cakici[2] and Andrea Ballestero[3] ⓘ

Abstract

Critical algorithm scholarship has demonstrated the difficulties of attributing accountability for the actions and effects of algorithmic systems. In this commentary, we argue that we cannot stop at denouncing the lack of accountability for algorithms and their effects but must engage the broader systems and distributed agencies that algorithmic systems exist within; including standards, regulations, technologies, and social relations. To this end, we explore accountability in "the Generated Detective," an algorithmically generated comic. Taking up the mantle of detectives ourselves, we investigate accountability in relation to this piece of experimental fiction. We problematize efforts to effect accountability through transparency by undertaking a simple operation: asking for permission to re-publish a set of the algorithmically selected and modified words and images which make the frames of the comic. Recounting this process, we demonstrate slippage between the "complication" of the algorithm and the obscurity of the legal and institutional structures in which it exists.

Keywords

Algorithms, normativity, accountability, responsibility, mystery, detective

This article is a part of special theme on Algorithmic Normativities. To see a full list of all articles in this special theme, please click here: https://journals.sagepub.com/page/bds/collections/algorithmic_normativities.

Algorithms as mysteries

Pursuing algorithmic accountability often translates to unravelling technical mysteries. As scholars and publics, we may be deeply fascinated by and personally invested in tracing clues and uncovering culprits. The promises are powerful: when an algorithm's authors can be identified, they can be held accountable for the effects of their work. In the context of normative and highly regulated legal systems, our tendency is to conceive of infractions against public wellbeing or common decency as punishable—if only we can reveal the answers to a few key questions.

The word "algorithm" has been used to refer to diverse sets of procedures, implemented in a variety of conditions with many potential outcomes. However, for all of us who cannot imagine algorithms as apolitical, the presence of an algorithmic operation indicates not rationalization but instead baked-in biases and unintended consequences (see Donovan et al., 2018). Whether defined as an idealized process or its implementation, an algorithm under investigation becomes a sort of smoking gun or bloody footprint, a trail not yet gone cold to trace as we search for further implications and responsible parties. The very presence of an algorithm, in other words, can be like the beginning of a detective story *in medias res*.

When we approach algorithmic operations as if they were mysteries, we treat the clues that they offer as evidence of something. Often that "something" is a process of events that relate, eventually, to authorship; someone or something that can be held accountable for

[1]Engineering, Design, & Society, Colorado School of Mines, Golden, CO, USA
[2]Technologies in Practice, IT University of Copenhagen, Copenhagen, Denmark
[3]Department of Anthropology, Rice University, Houston, TX, USA

Corresponding author:
Elizabeth Reddy, Engineering, Design, & Society, Colorado School of Mines, 1500 Illinois Street, Golden, CO 80401, USA.
Email: reddy@mines.edu

any ill effects. If the mystery is solved and that author revealed, justice can potentially be done. The innocent can be vindicated, victims given recompense, and the guilty may even be made accountable somehow—though perhaps not by way of the sort of moral drawing-room confrontation that Agatha Christie described.

Whether the ethical concerns that emerge around algorithms are explicitly epistemic, normative, or a matter of traceability (as in Mittelstadt et al., 2016), algorithmic governance is a problem that many understand to be pressing (see the review presented in Danaher et al., 2017). One response to this problem is to suggest that algorithmic processes should be made transparent exposed for supervision (Diakopoulos, 2013; Grimmelmann, 2008; Kraemer et al., 2010), revealing information that might otherwise require savvy detective work to uncover. Critics decry obscurity and suggest thoughtful legal and technical methods for limiting it (Crawford and Schultz, 2014; McKelvey, 2010; Pasquale, 2011, 2015). In this essay, we consider the pursuit of accountability within larger-scale normativities and address relationship between accountability and transparency in process and authorship.

The role of transparency and authorship in establishing algorithmic accountability is especially important to trouble. Transparency is a political technology that aims to make visible that which remains obscure (Ballestero, 2012). The very utility of algorithmic processes is that they remain black boxed as they travel further from the site and the moment of their production, and, as such, are mysteries. Even when algorithms have very overt effects, some of their technical details can be difficult for non-experts to parse (see, for example, Reddy, 2016)—indeed there is a point at which that black box becomes impossible for any human to penetrate. This being the case, public conversation about how exactly an algorithm was developed and used and who played pivotal roles might make an algorithm available for critique, focusing attention on particular events or more "subtle patterns of problematic behaviour" (Sandvig et al., 2014: 3). However, as Ananny and Crawford (2018) argue, even if algorithmic black boxes are opened up, that visibility may never be sufficient to bring about a connection between author, algorithm and consequences that is substantial enough that someone might be held accountable for an algorithm's undesirable effects.

However, it seems that a slippage in our understanding of accountability has sedimented in the popular imagination of algorithms. That slippage presumes that because many algorithms are constituted by obscure technical operations, the process for creating accountabilities implies remedying obscurity by finding out who are the authors as responsible parties. That is not necessarily the case. Accountability is not a matter

of discovering who or what is the author of a technical operation or information about how it was intended to perform, but a question about who or what should be made to answer for those operations. The question of accountability is caught up in normativities, shared upon values embedded in particular norms that legal and moral orders accept and protect. It cannot be reduced to authorship. Adopting an approach that builds uncritically on enlightenment epistemologies that assume that "*seeing* a phenomenon creates opportunities and obligations to make it accountable and thus to *change* it" (Ananny and Crawford, 2018: 2, emphasis theirs) is not sufficient.

When we frame algorithmic accountability as a product of transparency, particularly transparency with respect to process and authorship, we are neglecting the larger-scale normativities in which algorithms are embedded and on which accountability depends. Responsibility for the consequences of an algorithm do not necessarily need to be allocated to its author(s).

Making algorithmic processes and their implications more transparent is important work. We do not want to dismiss it. We want to think, however, about what we expect when we attempt to attribute responsibility by asking "who did it?"; by focusing our efforts on revealing how an algorithm works and who can take the blame for its effects. Tracing clues to reveal authorship and its processes might not be enough for accountability. The troublesomely attractive notion that one might productively "get down to" a truth of an algorithm (as Seaver, 2013 put it) runs the risk of unhelpfully reifying algorithms as distinct from other processes or formations (Neyland, 2016). We suggest that critical scholarship on algorithms could productively attend to the distributed conditions that algorithms exist within. Algorithms should be recognized as existing through other standards, regulations, and agents—these are the larger-scale normativities we alluded to above.

In this short commentary, we think about algorithms as mysteries to engage with conventional assumptions about how transparency can facilitate accountability when we learn the identity of the author of an algorithm, just as it might if we learned the identity of a murderer. We suggest that, although questions about accountability may be explored through attention to algorithmic authorship, a meaningful link between authorship and accountability only exists in the context of legal normativities. Whatever its promises, maybe all revealing authorship does is unveil the ways that authorship has been embedded in broader legal, economic, political, material, and/or moral systems. What we recognize as algorithmic practice and its products are embedded in conditions that include and exceed them (as in Reddy, 2018). This might not be a very satisfying conclusion for a mystery, but it does show

us how important it is to broaden our understanding of algorithms to include the normative contexts in which they exist. We demonstrate this by taking up the mantle of detectives ourselves and engaging in an algorithm-appropriate investigation of the authorship of—and through it, legal accountability for—an algorithmically generated comic. This comic, a provoking and playful art object called the Generated Detective, also makes use of the dramatic arc of a detective story genre to expose and challenge its conventions. We describe both the comic and our experience seeking permission to legally reproduce several of its frames to highlight the nuances of normative accountability that inform detective stories and algorithmic analysis alike. As we narrate the processes in which the Generated Detective is enmeshed, we ponder the conventions of legal accountability, including the technical and social issues that they foreground and those that they obscure.

On generating detectives

The Generated Detective[1] was created and released in a series of 15 short instalments in 2014 under the name Greg Borenstein. It was part of "National Novel Generation Month," or NaNoGenMo, a play on the annual internet-based "National Novel Writing Month" that writers organized first in the United States and then internationally since 1999 to spur their production in the month of November.

The pages of the Generated Detective are black and white. The 15 issues of the comic's arc hang together on the noir comic aesthetic that they share: stark images paired with sparse, strange, and driving text. These detective genre conventions signal readers to expect the kind of confusion and then slow progress toward clarity that are distinctive of detective stories, but the Generated Detective turns that promise of emerging order upside down. Here, words and images are borrowed from pre-existing works. The Generated Detective is, according to its subtitle, "An Algorithmic Comic" the way it was assembled both works within and subverts genre conventions.

Readers will not find a coherent plotline in the comic, much less a description of a mystery being solved. Instead of reading about a detective working to make sense of crime and reveal its perpetrators, here the reader becomes a detective with limited options. As they struggle to connect text to image, one page to the next, and to make sense of the comic's senselessness, readers themselves become detectives. The question of authorship that powers detective fiction—authorship and, of course, accountability—becomes a pursuit for the reader not *within* the pages, but *about* them.

It is much easier to make sense of how the Generated Detective was made than it is to discern a coherent plot thread within it. On the GitHub page for the project, Greg Borenstein describes the evolution of the software he used to produce the comic. The initial script searches a collection of detective novels legally designated to no longer be owned by their author, and instead, for all they might be attributable to one person or another, held in the public domain and available through Project Gutenberg.[2] The script returns sentences that match a given keyword from this database. Borenstein then chooses some of these sentences, searches image and video hosting site Flickr,[3] using the sentence as a search query to find a matching photo. A Processing[4] script gives the image the black and white comic book panel look. In later versions, he also searches sci-fi novels, modifies the placement of text boxes and image layout, experiments with using the same keyword for all panels, gathers sentences from genres instead of by matching keyword, and automates the selection of sentences from books and images (see GitHub, 2014).

The paragraph above describes a person called Borenstein authoring an algorithm that produces the Generated Detective. This we understand readily from the evidence in the comic itself, from how the comic circulates, and from conversations on GitHub discussion boards. The algorithm orders words and images, choosing among them as it works through its script. The authorship of the resulting comic, however, is not as easily understood, as it is distributed between Borenstein, the algorithm, and all the intermediate steps that make this production possible in a process that Ekbia and Nardi term *heteromation* rather than automation (2017). To human, automatic processes, and intermediate steps, then, we might attribute the position of writer and artist. Further authorship might be allocated to those who produced the words and images that Borenstein and his algorithm make use of. On GitHub, Borenstein celebrates the ambiguity of his role. He writes that maintaining some core authority allows him to make a comic that he finds interesting. He also celebrates the "accidents" and "flukes" of automation more than the successes of his own intentional interventions (GitHub, 2014).

If our interest in authorship were simply about determining how the comic had been brought into the world, we could conclude here. The technical procedures have been revealed, who and what did it is now understandable. However, we are interested in who we could consider to be formally and legally accountable for its effects. To replicate some of the comic's images, for example, we needed to consider the comic in the context of the system of copyright law. We also needed to establish who, in a legal sense, owned the panels of the comic featuring images and texts assembled from books, algorithmically processed and presented together.
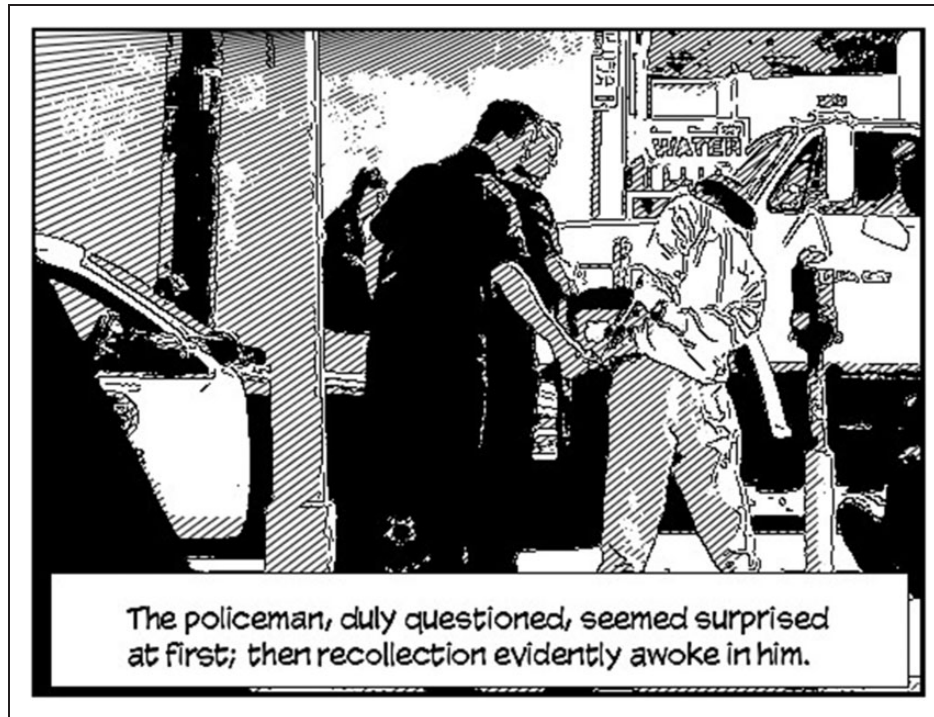
Figure 1. A panel from the first issue of Generated Detective.

We could, at this point, have continued our detective work to trace the texts' original authors. If some transparency was helpful, then why not try to reveal more? Legal norms prevented it. The texts' authors fell outside the scope of our inquiry. What they made is in the public domain through Project Gutenberg. Whoever penned them, inherited them, or bought them has no legal responsibility for them when it comes to issues of distribution and reproduction. A trail which might have, for example, taken us through a simple text search to Joseph Smith Fletcher's 1922 book *The Middle of Things,* or to publisher Alfred A. Knopf (as the text in both Figures 1 and 2 did), had to be abandoned. For our interest in copyright, it was a red herring.

The original images for the comic on Flickr presented us with a different kind of trouble, but we reached a similar conclusion. The processed versions with added text above were only available on the Generated Detective webpage, and the originals were drawn from a selection legally available for repurposing and significantly altered afterwards. While it may have been possible to trace them back to their sources, whether by reverse image lookup, or by consulting Borenstein, the search would not yield a solely responsible author.

Conscious of the form of accountability that mattered within copyright law pertinent to *Big Data & Society* and Sage Publishing, we disregarded extraneous leads and abandoned efforts to interrogate Flickr,

Fletcher's writing, Processing scripts and even musings on the production of the comic at GitHub. Many authorship functions collapsed into the one we had use for: Greg Borenstein himself. After a short email exchange, he granted his permission for us to replicate the images in this paper. The replication issue was resolved. We had our accountable-enough author.

An unsatisfying resolution

As we conducted this exercise, the relationship between transparency and accountability acquired new dimension. It was surprisingly easy for us to connect with Borenstein, to secure his authorization to use the frames in the story, and to project our relationship with him into the future by promising to give him a copy of our commentary. Tracing relations to the authors of the words and images drawn into Borenstein's story was a totally different matter. Those leads were dropped. Largely, we set these clues aside because there was a shared system of accountabilities—the copyright law pertinent to this publication—that spared us that work. Regardless of our inability to trace specific answers to questions like "who did it?" and "who is responsible?" beyond Borenstein, the legal system precluded the need to doing so. Little is at stake in our reproduction of the frames from the Generated Detective, and it may seem inappropriate to draw conclusions about algorithmic forms based on this simple and playful case. In our reading, however, the implications of our interactions
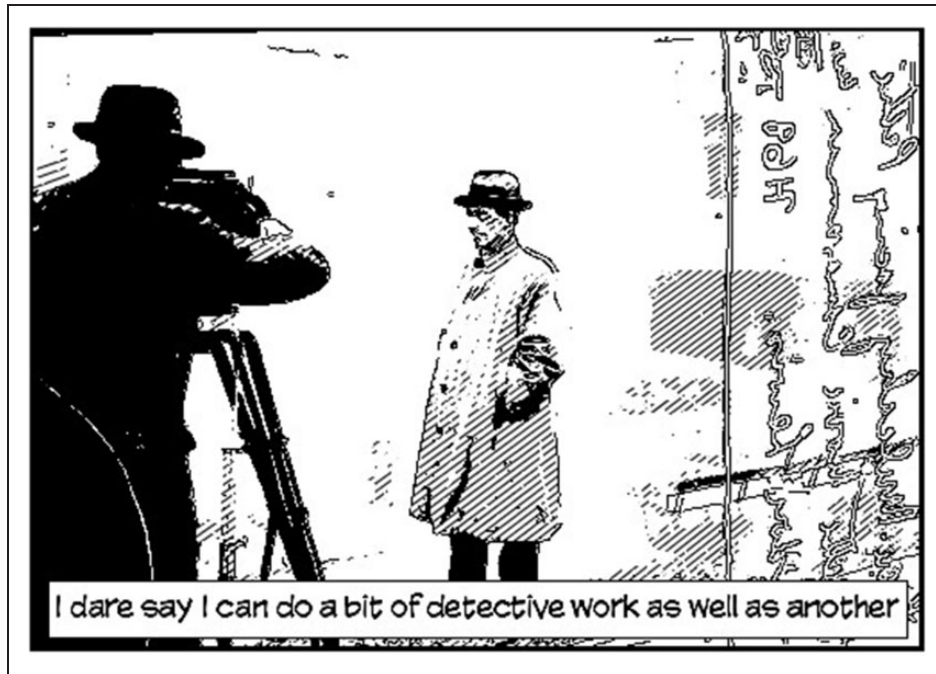
Figure 2. A panel from the second issue of Generated Detective.

with Borenstein are telling; Borenstein's willingness to become accountable and the copyright law system that allows him to do so, are both essential. Regardless of how distributed the authorship of the pieces that make up the story ultimately is, Borenstein has access to a larger-scale framework that allows him to step up to own Generated Detective.

We do not seek to propose a method for distributing accountability here; only to demonstrate how the transparency we achieved through detection is bound by legal accountability. Our broader point is that algorithmic accountability is complicated not because an algorithm is complicated, but because making an actor/author accountable is a complex issue that arranges large-scale agencies, social relations and histories in diverse ways depending on what is at stake. To build the chain of potential authorship, we needed to understand how the Generated Detective puts the pieces together, that is, to have a functioning understanding of how the algorithm does its work. However, at each step, our detective work demonstrated the entanglement of algorithmic work with other processes and agents. Revealing the algorithm and who stepped up as its author in ways that were sufficient for reproduction permissions did not resolve a mystery, in the detective story sense, but it did resolve our immediate problem. We secured an author and an authorization. In narrow regulatory terms, we had made somebody accountable for the story; in broader understandings of algorithms as social processes, we did not even begin to scratch the surface. Our work, like the comic itself, entailed no satisfying revelation.

This is the conundrum of accountability, particularly in the case of algorithms. Nothing intrinsic to the algorithm itself makes it easy or hard to decipher who is accountable for its consequences. Fully tracing the leads we had available became unnecessary when Borenstein was willing to claim authorship, grant authorization for reproduction, and continue the conversation. Even the small amount we were able to uncover with respect to the comic's complex authorship was well in excess of our legal needs. It was a simple enough process, though under different circumstances—a different algorithm with different financial consequences—the story would most likely have been radically different. This conundrum reflects broader political and historical decisions to protect, expose, or identify certain authors, and not others, in relation to diffuse values such as subjectivity, justice, responsibility, ethics, and capitalist accumulation.

Our playful case is an opportunity to reflect on the question of how transparency and accountability are created with respect to technical content. Revealing an author is not always the means to achieve accountability; having somebody step up to claim authorship is enough. In other words, we want to draw attention to how in the search for algorithmic accountability there has sometimes been a slippage between the "complication" of the algorithm as technical mystery with diffuse authors and the broader legal, ethical and institutional structures within which it exists and which determine who can be made accountable. It is not necessarily due to the technical machineries and

complicated chains of algorithmic forms working in connection with each other that accountability is difficult. Rather, it is difficult because it is hard to hold institutions and individuals accountable for the effects of their actions using legal norms that make protect parties that we would want to make responsible for the consequences of algorithms, regardless of whether they are the authors or not. Those protections are embedded in broad social and institutional structures.

Accountability is subject to convention. Who or what can be understood to responsible, and the consequences of their responsibility, are determined in relation to already established legal and ethical principles. The bureaucratic institutional form, the notion of private property, the idea of corporate personhood, and the legal bounds of juridical causality can inform our detective work but keep consequences beyond the reach of the accountability detective. An algorithm is never just a technical process. It requires integration into a broader sociotechnical system to work (see, for example, Eubanks, 2018; Noble, 2018; Schüll, 2012) and effective accountability measures must acknowledge that entanglement.

Put another way: the problem with a Facebook ad-targeting algorithm that facilitates the circulation of misinformation (Nadler et al., 2018; Vaidhyanathan, 2018) is not only that it cannot be deciphered, or even that the people who designed it cannot be revealed to the public. The problem with the way that Amazon Mechanical Turk integrates work management algorithms with the labour of low-status and easily exploited workers (Irani, 2015) is not only that the interfaces between automated and human processes are obscured.

We agree with the response that Sandvig et al. arrived at when they asked themselves: "'Is it useful to ask if "an algorithm" is unethical?' and 'Is "the algorithm" a useful focus of legal or ethical scrutiny?'" (2015: 4). It is absolutely essential to ask these questions. However, while understandings of authorship and accountability may predispose us to evaluate an algorithm's ethical effects and look for a culprit, the broader structures of corporate law protect Facebook and Amazon and the people within them from being effectively accountable for the consequences of the ways they shape social life. Embedded as they are in context and convention, there is little accountability to be demanded from the legal author, the Facebook or Amazon corporation, and virtually none from the coders who imagined those filters or processes. What would happen if we did not try to make somebody responsible for the effects of an algorithm by tracking who is its author; if we considered the real complexities of accountability in and through both? What we argue is that, instead, we could focus on how allocating responsibility is a matter of convention, that is,

a matter determined by a broader set of legal, ethical and political principles, a technolegal metaphysics (Ballestero, 2015), that may have an algorithmic body but embodies much more than code.

As detectives, we have uncovered something of a chain of authorship for the Generated Detective that is adequate to our immediate needs. We have followed clues, uncovered some leads and abandoned others. In the process we have revealed the degree to which certain mysteries around algorithms may be, simply, easy to solve, and that is the trouble we expose here. The conventions of detective fiction promise that tracing clues to uncover the workings of an algorithm will help us allocate responsibility and make remedies possible, but our investigations must also consider norms of accountability itself. That is, instead of solely pursuing the question of who did it, we may reach more productive resolutions by interrogating the structures in which accountability is embedded. Taking algorithmic accountability seriously must mean considering the legal and social norms that structure its conditions of possibility. In our critiques and intervention, we need to remember what the darkest detective fiction reveals: finding out who did it, where, and with what instrument matters, but the answer does not bring the culprit to justice on its own.

Notes

1. Available in its entirety through http://gregborenstein. com/comics/generated_detective/
2. https://www.gutenberg.org/
3. https://www.flickr.com/
4. https://processing.org/

ORCID iD

Elizabeth Reddy ⓘ http://orcid.org/0000-0003-0673-1173
Andrea Ballestero ⓘ http://orcid.org/0000-0001-7692-0514

References

Ananny M and Crawford K (2018) Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media and Society* 20(3): 973–989.

Ballestero A (2012) Transparency in triads. *PoLAR: Political and Legal Anthropology Review* 35(2): 160–166.

Ballestero A (2015) The ethics of a formula: Calculating a financial humanitarian price for water. *American Ethnologist* 42(1): 262–278.

Borenstein G (2014) Generated Detective. Available at: http://gregborenstein.com/comics/generated_detective/ (accessed 27 February 2018).

Crawford K and Schultz J (2014) Big Data and due process: Toward a framework to redress predictive privacy harms. *Boston College Law Review* 55(1): 93.

Danaher J, Hogan MJ, Noone C, et al. (2017) Algorithmic governance: Developing a research agenda through the power of collective intelligence. *Big Data & Society* 4(2): 1–21.

Diakopoulos N (2013) *Algorithmic Accountability Reporting: On the Investigation of Black Boxes. A Tow/Knight Brief*. New York, NY: Columbia Journalism School, Tow Center for Digital Journalism.

Donovan J, Caplan R, Matthews J, et al. (2018) Algorithmic accountability: A primer. *Data & Society Tech Algorithm Briefing: How Algorithms Perpetuate Racial Bias and Inequality*. Washington, DC, USAAvailable at: https://datasociety.net/wp-content/uploads/2018/04/Data_Society_Algorithmic_Accountability_Primer_FINAL.pdf (accessed 11 November 2018).

Ekbia HR and Nardi BA (2017) *Heteromation, and Other Stories of Computing and Capitalism*. Cambridge, MA: MIT Press.

Eubanks V (2018) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York, NY: St. Martin's Press.

GitHub (2014) Issues thread: Generated detective: A NaNoGenMo comic. Available at: https://github.com/dariusk/NaNoGenMo-2014/issues/70 (accessed 27 February 2018).

Grimmelmann J (2008) The Google dilemma. *New York Law School Law Review* 53: 939.

Irani L (2015) Difference and dependence among digital workers: The case of Amazon Mechanical Turk. *South Atlantic Quarterly* 114(1): 225–234.

Kraemer F, Overveld K and Peterson M (2010) Is there an ethics of algorithms? *Ethics and Information Technology* 13(3): 251–260.

McKelvey F (2010) Ends and ways: The algorithmic politics of network neutrality. *Global Media Journal: Canadian Edition* 3(1): 51–73.

Mittelstadt BD, Allo P, Taddeo M, et al. (2016) The ethics of algorithms: Mapping the debate. *Big Data & Society* 3(2): 1–21.

Nadler A, Crain M and Donovan J (2018) Weaponizing the digital influence machine: The political perils of Online Ad Tech. *Data & Society* https://datasociety.net/output/weaponizing-the-digital-influence-machine/ (accessed 14 November 2018).

Neyland D (2016) Bearing account-able witness to the ethical algorithmic system. *Science, Technology, and Human Values* 41(1): 50–76.

Noble SU (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

Pasquale F (2011) Restoring transparency to automated authority. *Journal on Telecommunications and High Technology Law* 9: 235–254.

Pasquale F (2015) *The Black Box Society: The Secret Algorithms that Control Money and Information*. Cambridge, MA: Harvard University Press.

Reddy E (2016) The production of earthquake emergencies: Technoscientific earthquake early warning in a social and material world. *Revista Iberoamericana de Comunicación* 30: 17–44.

Reddy E (2018) 'A world we don't know': The spatial configuration of sensory practices and production of knowledge in and around Mexican seismic monitoring. *Journal of Political Ecology* 25(1): 509–524.

Sandvig C, Hamilton K, Karahalios K, et al. (2014) Data and discrimination: Converting critical concerns into productive inquiry. *Paper presented to Data and Discrimination: Converting Critical Concerns into Productive Inquiry at the Annual Meeting of the International Communication Association* Seattle, WA, USA. Available at: https://pdfs.semanticscholar.org/b722/7cbd34766655dea10-d0437ab10df3a127396.pdf (accessed 11 November 2018).

Sandvig C, Hamilton K, Karahalios K, et al. (2015) *Can an Algorithm be Unethical Annual Meeting of the International Communication Association*. San Juan, PR, USA.

Seaver N (2013) *Knowing algorithms*. Paper presented at Media in Transition 8. Cambridge, MA, USAAvailable at: http://nickseaver.net/papers/seaverMiT8.pdf (accessed 27 February 2018).

Schüll N (2012) *Addiction by Design: Machine Gambling in Las Vegas*. New Haven, NJ: Princeton University Press.

Vaidhyanathan S (2018) *Anti-Social Media: How Facebook Disconnects Us and Undermines Democracy*. New York, NY: Oxford University Press.