Coded Governance

By
Nick Lally

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy
(Geography)

at the
UNIVERSITY OF WISCONSIN–MADISON
2018

Date of final oral examination: 8/2/2018

The dissertation is approved by the following members of the Final Oral Committee:
Keith Woodward, Associate Professor, Geography
Sarah Moore, Associate Professor, Geography
Lucas Graves, Associate Professor, Journalism and Mass Communication
Robert Roth, Associate Professor, Geography
Luke Bergmann, Associate Professor, Geography, University of British Columbia

Abstract

In this dissertation, I describe how existing practices of governance become integrated with software, shifting how spaces are constructed, understood, and governed. The term "coded governance" here denotes the move towards algorithmic modes of decision-making, judgment, surveillance, and policing. "Coded" in this context refers to (1) computer code on which software relies, (2) the proliferation of opaque systems whose inner workings are hidden from public view, and (3) the automated classification of people and places made possible by software. Building on earlier mapping and statistical methods, software systems used for governance expand the possibilities for identifying patterns across disparate datasets, tracking and surveilling individuals, automating decision-making processes, and making predictions about the future (Amoore 2013). How software becomes unevenly integrated into infrastructures of governance is the subject of the three case studies that make up this dissertation. Chapter 2 focuses on how algorithms change existing, non-digital infrastructures, making them reconfigurable, interoperable, and deceptive. The unique characteristics of algorithmic infrastructures, I argue, have important implications for the coding of people and places. Chapter 3 describes an online effort to crowdsource the investigation of the 2013 Boston Marathon bombing. I show how disparate systems became interlinked as internet users revealed new possibilities for surveillance and intervened in the event as it unfolded. Finally, in Chapter 4, I draw from interviews with developers of predictive policing software in the US and UK to show how they navigate the technical, ethical, and practical challenges of producing software for policing. I show how uncertainty arises during key moments in the process of modeling crime data, making predictions about the future, and integrating software into police departments. Informed by the three case studies, the conclusion includes three methodological suggestions for the the field of digital geographies, which suggest creative and exploratory orientations towards software as an object of study. Drawing from these suggestions, I conclude with a proposed research program for the future study of predictive policing software.

Acknowledgments

I am grateful to so many people for supporting this work at various stages of its development. More broadly, I am happy to have found a home in geography and appreciate all of the people who have welcomed me into the discipline, challenged me intellectually, and made this all possible.

First, thanks to my committee members, whose feedback and support throughout this process have been instrumental in shaping these ideas: Keith Woodward, Sarah Moore, Lucas Graves, Rob Roth, and Luke Bergmann. I am grateful to so many others have engaged with my work, given feedback, and contributed to my scholarly development along the way, including: Bob Kaiser, Erika Marin-Spiotta, Kris Olds, Jim Thatcher, David O'Sullivan, Liz Hennessy, Kristin Eschenfelder, Samer Alatout, Lyn Macgregor, Sai Suryanarayanan, and Matt Wilson. To everyone in the UW Cartography Lab, thank you for all of the lunches, discussions, and support: Tanya Buckingham, Meghan Kelly, Daniel Huffman, Leanne Abraham, Ross Thorn, and Alicia Iverson. Thanks to little group for being excellent co-advisors and co-conspirators: Elsa Noterman, Charlie Carlin, Jessi Lehman, Luke Leavitt, Doug Adams, and Mario Bruzzone.

I cannot imagine a better place to have done this work than the Department of Geography at the University of Wisconsin–Madison. I am appreciative to all of the people who have contributed to and continue to build the collaborative, supportive, and vibrant academic culture of the department, including (but certainly not limited to): Eric Nost, Danya Al-Saleh, Rafi Arefin, Kelly Chen, Nathan Green, Niwaeli Kimambo, Rebecca Summer, Ruth Trumble, Elliot Vaughan, Nicolle Etchart, Carl Sack, Marguerite Roulet, Morgan Robertson, and Lisa Naughton. And finally, thanks to Kallista Bley for sticking with me through this journey and always encouraging me along the way. I look forward to many more adventures!

This research was made possible by generous funding from the Holtz Center for Science

& Technology Studies, the Scott Kloeck-Jenson fund, the Department of Geography, and the Graduate School.

Table of contents

1. Introduction

Beginning in 1886, researcher Charles Booth set out to understand poverty in London.[1] Accompanying what would eventually become a seventeen volume study were maps showing the spatial distribution of poverty at the block level. In the original poverty maps released in 1889, Booth developed seven categories to represent social classes, which were used to color-code the streets of London. At the top of the class spectrum, represented in yellow, were the "Upper-middle and Upper classes, Wealthy," while descending social classes were colored red, pink, purple, light blue, blue, and finally black. Most of the classifications focused on earnings and needs to make class distinctions, which included phrases like "Good ordinary earnings," "18s. to 21s. a week," and "Chronic want". However, at the bottom of the class hierarchy we find the "Lowest class. Vicious, semi-criminal" colored black on the map ("Charles Booth's London" 2018).

Booth's maps use economic status as a proxy for criminality, reflecting a particular orientation toward the world, and it is not insignificant that much of the data was collected on walks with police on patrol. In the field notes of Booth's researchers from 1899, for example, a police inspector, when asked about the people inhabiting Chadwick Street, "appeared to think that thieving and prostitution were the chief occupation of the people," agreeing that the block should be colored black on the map (Arkell and Aves 1899).[2] While Booth claims an objective orientation towards his subjects in which he merely translates facts into quantitative measures and representational codes (Booth 1889), his maps produce spatially-bounded subjects who

[1] The connection between Booth's maps and maps used for policing, as explicated below, is drawn from Andrew Evans's lecture titled "Crime prediction: recent approaches" given at Maynooth University on February 15th, 2018. Copies of Booth's maps, it turns out, were framed and displayed in one of my field sites: a software company that produces and sells predictive policing systems.

become knowable and governable through the god's eye view of the map (Haraway 1988). The limited categorization scheme of the map, which makes a visual argument demarcating who does and who does not inhabit city blocks, inscribes difference onto the city. If the police transform individuals into subjects by hailing them ("Hey you!"), this process of ideological interpellation (Althusser 2001) is concretized in the map, which can then be called upon to inform future processes of governance.

How maps become enrolled within political processes depends not only on their representational schema, but also on the positions they occupy within socio-technical networks (Kitchin and Dodge 2007). In the case of the Booth maps, they could be used to inform progressive social change—the allocation of social services to needy communities, for example—just as easily as they could be used to target policing towards the "vicious, semi-criminal" subjects who inhabit the black drawn streets of London ("Charles Booth's London" 2018). While Booth's intent was more aligned with former, there is no secure epistemological status of the representational schema of the map once integrated into existing socio-technical networks. The maps, then, reflect an ideological position in relation to social and spatial relations while informing differential possibilities for governing and remaking the space of the city.

The power of maps to inform political decision-making and remake the world to reflect particular ideological orientations is a story at least as old as cartography itself (Harley 1989). Supported by maps arguing that the cramped quarters of downtown Paris bred crime and disease, Georges-Eugène Haussmann with the support of Napolean III reimagined a modernized Paris with grand boulevards that cut through the tiny, meandering, and ungovernable streets of the city center (Picon 2003). The new, wide boulevards would make visible the previously hidden nooks and crannies of the complex street networks while making impossible the barricades that were strategically deployed during uprisings (Benjamin 1969). The results of these efforts transformed the spaces of Paris and with them, the possibilities for both urban governance and political resistance. Paris was transformed into a city designed and imagined for and by the bourgeoisie who could better control and manage class divisions.

Similarly, the social and spatial structuring of nearly every city in the United States has been deeply influenced by the power of maps. Beginning in the 1930s, the Home Owners' Loan Corporation (HOLC) produced "residential security maps" for cities across the country, which color-coded neighborhoods according to their assessed real estate credit risk. These redlining maps, rather unsurprisingly, reflected existing patterns of racial and economic segregation (Hillier 2003), further entrenching the uneven geographies of American cities. They were then both a reflection of racist spatial imaginaries, but also worked to further sediment existing inequalities built over decades of racist policies, beliefs, and everyday practices (Woods 2012). The maps, then, reflect how racism can be understood as a spatial problem, built into the structure of the city, which continues to weigh upon the present. While homeowners in the affluent and largely white neighborhoods that were deemed to be low risk for home loans built real estate equity, those in the high risk areas were denied similar possibilities for intergenerational wealth accumulation (Aaronson et al. 2017).

If nineteenth and twentieth century practices of governance and political decision-making were both structured by and reflected in static maps that represented data, the more recent shift to computational approaches to collecting, sorting, analyzing, and representing spatial data has open new possibilities for data to inform political decisions. For some, big data, machine learning, and artificial intelligence promise to transform what we are able to know about the world, revealing new patterns that reflect a dynamic but knowable world. If these boosters, at their extreme, mark this shift as the "end of theory" as theories can be extracted from existing data (Anderson 2008), others have shown how automated, data-driven approaches merely reify and entrench existing inequalities, often under the guise of scientific objectivity (Eubanks 2017; O'Neil 2016). In response, geographers have called for critical approaches to studying the data (Dalton, Taylor, and Thatcher 2016), algorithms (Kwan 2016; Kitchin 2017a), technologies (Leszczynski and Elwood 2015), and ideologies (O'Sullivan 2018) that make the production, analysis, and use of big data possible. Instead of a value-neutral practice, these scholars outline the decisions, contingencies, contexts, and assumptions that structure the insights produced through big data.

Not only are the claims that big data can make about the world fraught and contested, they also open possibilities for intensified personal targeting and surveillance. Individual subjects are constructed and governed through inferences made from the constant data streams produced by their interactions with cell phones, sensors, surveillance systems, the internet, institutions, and other systems that have become embedded within the spaces of everyday life (Christl 2017). Big data makes possible both the categorization and governance of aggregated spaces, but also of individuals, both of which can be measured, analyzed, and modulated through automated processes (Cheney-Lippold 2011). These processes of individualized targeting and surveillance often reflect and serve to further entrench existing racial, social, and economic inequalities (Browne 2015). How subjects and spaces are constructed and understood has important implications for systems of governance that increasingly rely on automated processes.

This dissertation takes as its starting point the shift to computational and algorithmic modes of governance, tracing these practices as both a continuation of earlier mapping and statistical methods, but also showing how computation changes how the world is constructed, understood, and governed. Dynamic mapping, machine learning, online data streams, and computational infrastructures are the technological means through which computation, working within socio-technical networks, comes to reshape spaces of everyday life. With the growing interest in digital geographies in the discipline of geography comes a proliferation of theories and methodological orientations towards the spatial implications of computation (Ash, Kitchin, and Leszczynski 2018). Building on a long history of engagement with computation and cartography, geography is well-positioned to not only study, but also intervene in these processes of world building. Many of the existing problems with the map—problems of categorization, ideology, and governance like we saw in the Booth maps—persist in the move towards automation and real time analysis. But these systems also beget new problems and spatial imaginaries as they become integrated into infrastructures of governance, informing political decision making with little oversight or regulation. These new infrastructures, however, do not arise from nowhere, but are built upon existing knowledges and practices, many originating within the discipline of geography.

Problems with the map

The discipline of geography is deeply implicated in the practices of cartography, geographic information science (GIS), and spatial data on which current modes of algorithmic governance rely.[3] So too has geography offered a long history of critical orientations towards computation, which continue to expand with the growing interest in digital geographies. Just like grandiose claims around big data are met with critical voices that argue how existing forms of power are maintained or accentuated by computation, geography has experienced similar rifts that inform its current orientation towards the digital.

Beginning in the 1940s computation promised a transformation of the discipline of geography as spatial laws could be modeled and used to describe social behavior. This

[3] Many of the spatial analysts and computer programmers I have interviewed over the course of my research have been trained in or have connections to geography departments.

epistemological shift, from the largely descriptive regionalism of the time, to the law-seeking quantitative revolution, mirrors the current big data revolution, which similarly seeks to reveal the hidden patterns and laws that undergird the world. As the "space cadets" of geography took up the call of quantification, enabled in part by the punch card computers that were appearing on college campuses, discontent about the epistemological claims of the movement slowly began to grow (Barnes 2004). For some geographers, the grandiose (and often masculinist) claims of the movement necessitated a break with quantification, resulting in many geographers turning to marxism (e.g., Harvey [1973] 1988), humanism (e.g., Buttimer 1976; Tuan 1976), and feminism (e.g., Monk and Hanson 1982; Foord and Gregson 1986) in the 1970s and 80s while denouncing the allegedly positivist methods of quantitative approaches. There were, to be fair, scholars who bridged the divide, and whose work continues to inform current approaches that reject hard borders between critical and quantitative methods (see, for example: Bunge 1960; Tobler 1961; Gatrell 1983). But they were in the minority, especially as the history of the discipline is narrated, and louder voices prevailed, even if they themselves were deeply influenced by quantitative methodologies.

This intellectual schism would resurface again in the 1990s after GIS had grown to prominence within the discipline. In the GIS wars that followed, critical geographers would once again take the practitioners of quantitative methods to task, claiming that positivist methods had snuck back into the discipline through GIS. In a meeting at Friday Harbor in Washington in 1993, the two camps would meet to hash things out (Poiker 1995). While some progress would be made in integrating the two sides, including Initiative 19 of the National Center for Geographic Analysis (NCGS) calling for the study of the social implications of GIS, mainstream GIS practitioners initially withdrew from the discussion (Schuurman 2000). In spite of the largely one-sided engagement, quantitative approaches would not be marginalized as they had been following the quantitative revolution. Instead, GIS would assume a prominent role within the discipline as geography departments offered GIS training both as a research method and marketable skill.

For some US geography departments, what followed the GIS wars of the 1990s were uneasy truces within departments (Schuurman 2000). Despite working in the same spaces, this meant that human geographers and GIScientists occupied non-overlapping methodological and epistemological worlds. For others, however, critique offered an entryway into using and changing GIS as a critical tool, while GIS would become more open to reflexive and critical approaches. In the wake of Friday Harbor, a generation of scholars slowly emerged who could fluently travel between the technical world of GIS and the critical world of human geography, tacking back and forth between the two to produce others ways of knowing the world through computational means. Instead of an external critique, this group of scholars began developing constructive critiques that cared for its subject (Schuurman and Pratt 2002), paving the way for qualitative (Cope and Elwood 2009), feminist (Kwan 2002), and critical (Thatcher et al. 2016) orientations towards GIS. Using these new approaches, critical scholars would both use GIS as a research tool while recognizing its latent assumptions and limitations.

Reconfiguring the map

If critical GIS and cartography scholars developed reflexive tools that recognized the hidden power relations of the map, the recent move towards automated maps, predictive analytics, and real time data brings new theoretical and methodological challenges. The ability for software to dynamically respond to changing conditions in real time affords the possibility for automatic and continuous modulations of infrastructures. For example, as described by Rob Kitchin (2017a), data from sensors can be used to measure traffic flows, which can then be analyzed computationally, triggering traffic lights to manage and modulate traffic to maximize efficiency. These systems, Kitchin observes, do not only react to current conditions, but also deploy anticipatory logics that attempt to predict what will happen next and adjust the present conditions accordingly. Many computational systems purport to make the future calculable through the analysis of data streams, infusing city infrastructures with automated logics that attempt to preempt calculated futures (Anderson 2010; Amoore and Raley 2017). Since real time flows of data are always changing, the futures conjured by algorithmic logics are also always shifting, especially as machine learning allows computational systems to constantly adapt their models to fit new data.

Maps, following Kitchin et al. (2013), are here understood as processual objects, always unfolding through practices, contexts, design, and use, but also constantly shifting algorithmically. Maps emerge in relation to practice, immersed within different socio-technical networks that give them different meanings and imbue them with different possibilities for modulating space (Kitchin and Dodge 2007). For example, in Chapter 4 (below), the maps of predictive policing are taken up and understood in vastly different ways by various people, which includes developers, police officers, academics, and the media. People ascribe meaning differently to the categories of the map, rooted in their own situated orientations towards the city and crime. The categories of difference produced through the maps are taken up unevenly, producing different implications for understanding and governing the city. Additionally, different interpretations feed back into the production of maps as actors negotiate changes to the computational functions that produce predictive policing maps.

In studying systems like predictive policing, there are a number of approaches that can both complement and build upon lessons from critical GIS, as real time spatial analysis often builds upon methods that have long been a part of GIS software. At the center of these systems are computational processes that are enacted by software, which act upon data. Software encodes algorithms—described in computer science as abstract descriptions of steps used to transform data, often to solve specific problems—which then directs how a computer functions and acts upon data. How software is built and what algorithms it incorporates changes how software is able to process data to make claims about the world. On the one hand, a technical analysis of software can help reveal how these systems function, giving insights into their possibilities and limitations. Algorithmic audits, technical documents, and interviews with programmers can all contribute to building knowledge about technical systems. On the other hand, software only works in situated contexts, built through particular spatial imaginaries, working in particular places, situated within complex and shifting networks of relations, and unevenly modulating and producing the spaces of everyday life. Therefore, the study of mapping software requires an approach that attends to not just software as such, but also the relations and materialities that connect to it in every direction.

Placing automated software processes at the center of this study forces a rethinking of the place of the map within the socio-technical networks. The communicational model of cartography still weighs heavy as maps are the interfaces between the computer and the people who read them in order to form or validate political orientations and make decisions. The proliferation of interactive maps in particular requires additional attention to ascertain how they work in different contexts and configurations—a project that builds on existing understandings of cartography to open new questions around interactive maps (Roth 2013). User studies of interactive maps, for example, are beginning to reveal how representational and functional choices made in the creation of maps for decision makers will affect how the world is understood and how decisions are made (Vincent et al. 2018). However, the interactive or static map only offers a fleeting glimpse of automated computational processes that analyze spatial data.

Analogous to the map as a diagram of spatial relations, software relies on and, in the case of machine learning, is constantly building and remaking its own diagrams of spatial relations. These diagrams are only legible to the computer, relying on complex abductive logics to recognize patterns in higher dimensional spaces and make models to interpolate new data. For example, some predictive policing software will analyze distances between a number of geographic features and crime to extract patterns that can predict and map future crimes. The maps that result only offer a limited and partial view of the computational functions that produce them.

Just as Louise Amoore (2016) warns of mistaking NSA data centers for central sites of algorithmic power, so too are maps only the most visible manifestations of hidden processes of calculation. Following Amoore's suggestion, I look to the offices where algorithms are produced, validated, and reworked as well as in the functioning of code to ask how software informs practices of mapping and shapes spatial understandings. Interactive maps produced to inform policing, for example, are intermediary objects that translate automated practices of spatial analysis—which themselves are produced through fraught social negotiations—into visual representations for human interpretation. How these maps go out into the world and are acted upon is itself a contingent upon the socio-technical networks, relations, and practices in which they are immersed. The map, then, is only a starting point for understanding how political decision-making and everyday practices are being reshaped by algorithmic practices.

Theoretical Framework

In this study, I draw from science and technology studies (STS), software studies, and digital geographies to inform my methodological and theoretical approaches. Influenced by STS, I examine how claims about the world are produced within the networks of which algorithms are a part. Produced through negotiations, frictions, and contingencies, I look to the fraught processes of production and implementation that proceed the solidification of software into an infrastructural technology. Drawing from software studies, I am attentive to the specific computational processes enacted by algorithms with a focus on how they enable particular truth claims to be made. Additionally, both STS and software studies contribute to understanding how subjects are produced in relation to software. This includes those who work in collaboration with machines as programmers as well as those who become the subjects of algorithmic surveillance, policing, and other technologies of governance. Finally, the emerging field of digital geographies grounds STS and software studies, developing ways to understand how spaces are produced, understood, and modulated by and through software.

Science and technology studies

The ability for algorithms to extract patterns from vast amounts of data, using abductive reasoning with little regard to causality, and making certain phenomena perceptible and actionable, is the central political problem posed by algorithms according to Louise Amoore (2016). Drawing on Karen Barad's (2007) work on agential realism, Amoore argues that apparatuses of computing draw boundaries around and makes legible particular subjects, making them matter—adding people to no-fly lists and preventing passage across borders, for example. And the experimental nature of programming means that the programmer can experiment with thresholds that control pattern recognition, tweaking them to add more or fewer people to these lists. Amoore contrasts this against geographical accounts that attempt to understand the politics of networked computation by making legible data centers and cables—the physical things that we commonly think of as infrastructure. But, to extend Barad's work from which Amoore draws, I want to argue that both are important aspects to understanding the political possibilities of computation. As Barad (2007) argues, it is through particular material configurations of apparatuses, people, and other materialities that phenomena, identities, and subjectivities are brought into existence.[4] So understanding how we come to know the world through computation requires studying not just the symbol-manipulation procedures of particular algorithms, but the entire field of practices, social systems, hardware, people, and apparatuses that coalesce to afford a computer to take a cut on reality.

How a socio-technical field coalesces to produce knowledge is a central question in STS, of which Barad is a part. STS undermined the assumed objectivity of science by questioning the epistemological foundation of scientific knowledge, opening scientific practices to new forms of theoretical inquiry (see, for example: Bachelard 2002; Fleck 2008; Kuhn [1962] 2012). Arguing that the construction of scientific facts is the result of complex negotiations, ruptures, contingencies, and creative interpretations, all situated within particular social and material contexts, some scholars moved to the laboratory to understand how science is produced. In an early laboratory study, Latour and Woolgar (1986) found a chaotic scene of disordered observation through which scientists struggled to make a coherent and logical account of the world. Denying a relativist or a realist position, but positioning themselves somewhere in the middle, Latour and Woolgar argue that a fact is produced within a specific social network of practice with the use of particular devices and "cannot simply jump out of the very network of social practice which makes possible its existence" (183).

In later works, both Latour and Woolgar describe how subjects are produced and depend on networks of practice, which will be reflected in many approaches to actor-network theory (ANT). Woolgar (1990), for example, rejects the "ontological gerrymandering" that falsely orders

[4] To illustrate this argument beyond its origins in particle physics, Barad uses the example of a jute mill in Calcutta.

the world under "the presumption that entities are bounded" (63). In an ethnography of a computer firm, Woolgar focuses on the emergence of subjects through the process of human-computer interaction as "the evolving machine attempts to configure the user" (61). John Law (1992) similarly argues against thinking of a person as a bounded entity. Instead, he claims, "what counts as a person is an *effect generated by a network of heterogeneous, interacting, materials*" (383, emphasis in original). Latour's (2005) theory of the subject similarly relies on networks that are constantly in flux. Following Whitehead, he explains, "subjectivity is not a property of human souls but of the gathering itself—provided it lasts of course" (218). Here, Latour adds an element of temporality to the conception of the subject, as the gathering of a network is not guaranteed to last. As Roy Boyne (2001) explains, "Latour's view, essentially, is that the subject is eternally and irreducibly constructible... The human subject is inevitably and permanently hybrid. This does not amount, for Latour, to the death of the human, but it does mean that the human is not a stable form" (29). The human subject, then, is as unstable as the network that defines it. And the network, we learn, only exists in a continually state of making and remaking (Latour 2005).

Similarly, the epistemological promises of algorithmic modes of governance are rooted in specific social and material networks that produce insights about the world while constructing subjects. For example, in detailing the history of climate data and modeling, Paul Edwards (2010) shows how we came to know global climate change through political negotiations, scientific networks, power struggles, standardization and friction, and the constant reconfiguration of scientific knowledge. It is an example of how the world is understood in particular ways through complex and ongoing negotiations with data, people, and machines. Scientific negotiations and contingencies become forgotten or erased when a new scientific paradigm or fact is widely adopted (Fleck 2008; Kuhn [1962] 2012), but it is within those contested moments that STS approaches can reveal the complex politics of scientific production. STS, then, adopts both an historical and critical approach to understanding how science developed the way it did, with attention to the foreclosed possibilities that have been brushed aside. Algorithmic forms of governance are produced through similar moments of negotiation and uncertainty as new models attempt to make previously unseen phenomena visible through complex socio-technical assemblages coalescing around computing.

The negotiations that go into creating technical objects is an important problem for scholars, one that reveals not only the politics that have become sedimented within technologies, but also hinting at possibilities that have been foreclosed (Bijker 1995). Embedded within software, then, are the complex and often contested politics that went into its creation and shape its continued existence. As Kate Crawford (2016) observes in her examination of the agonistic politics of algorithms, "workplaces are themselves spaces of everyday conflict and contestation, where algorithmic design decisions are made after debate, disagreement, tests, and failures" (89). So long before an algorithm enters into use, it is produced in and through organizational arrangements that battle over its creation. It is in these moments—in the company office, during meetings with clients, and in public demonstrations—before software systems become solidified, that I look to examine the translations, negotiations, misunderstandings, and contested politics that later become erased, but which are central to the production of software.

Software Studies and Digital Geographies

The field of software studies, which has informed some approaches to the growing field of digital geographies (Ash, Kitchin, and Leszczynski 2018), gives insights into the political, social, and cultural effects of software (Wardrip-Fruin 2012). Geographers have shown how software has come to produce and reaffirm urban inequality (Graham 2005), make cultural objects endlessly mutable (Rose 2016), inform practices of environmental management (Nost 2015), mediate and construct language (Thornton 2017), and classify urban neighborhoods (Payne 2017). Central to these accounts are analyses of how computation produces and changes space in different contexts. With an increased interest in computation as an object of study within geography, scholars have increasingly turned to specific algorithms as matters of concern.

"Algorithm" has become a buzzword in the social sciences and popular press, often used loosely as a way to indicate something a computer has done or could do. For the purposes of this study, however, I refer to algorithms as they are defined in computer science. That is, as an abstract description of a set of procedures and calculations used to process and transform data, often as a means to solve a specific problem (Gillespie 2010, 167; Wardrip-Fruin 2012, 17; Striphas 2015, 403). This definition means algorithms are independent of their encoding in a particular programming language, instead, they indicate the processes that a computer will instantiate when code is run. The translation of the abstract language of algorithms into software is a messy process as it brings numerous languages, libraries, functions, and code bases together with unpredictable results, making the programming process one of experimentation and play. As a result, so complex are many software packages that "it is unlikely that any one programmer has a complete understanding of a system" (Kitchin 2017b, 21).

If software exceeds the ability of any one person to understand it completely, especially as algorithms build on each other, then calls for "algorithmic transparency" as a cure for bias are nearly impossible to implement (Crawford 2016, 87). But a focus on processes gives us some insight into how things work, without having to understand the code of a system as a whole. Computational processes can be recovered from historical and technical documentation that might give insight into the political debates that went into deciding on a particular set of algorithms; through analogy with other systems—most machine learning techniques, for example, are well-known and deployed widely; and through interviews with programmers who can describe the general functioning of complex software packages (see also: Kitchin 2017b). An understanding of algorithms as processes and as the result of political contestation places them within their wider political and social contexts. Posing them as an abstract description of processes—often likened to a recipe—becomes especially important in this study owing to the availability of general descriptions of many of the algorithmic methods that I write about in the chapters that follow.

How algorithms process data to make phenomena visible, and how we interpret those results, goes hand in hand with how we come to understand the world (Kirsch 1995; Dourish and Bell 2007; Mattern 2015). The rationalization, mechanization, and mathematization of the world, which we can trace back to the philosophy of Descartes ([1637]1984), is a foundational ideology for many attempts to model and govern the world using computation. As Joseph Weizenbaum (1976) argues, the "remaking of the world in the image of the computer started long before there were any electronic computers," but computers better enable us to see this transformation (ix). For some, computers became the tool that could finally model the precise mechanics of the world, producing what N. Katherine Hayles (2005) has called the belief in a "Computational Universe." Understandings of the world, in this account, reflects understandings of computers as both can be reduced to mathematical logics and data structures. The reliance on computers can also lead us to make over the world to make it computable. In the simple act of building numbering, Reuben Rose-Redwood (2012) shows how "representation has the capacity to reshape the world in its own image" as the space of the city is shaped by the need to fit into computational databases (299). Computing comes to inform the spatial imaginations of decision makers, affecting how the world is governed, and so it becomes important to understand the ideological underpinnings of computational models.

Not only a way of seeing the world, computation also allows for the extension of human agency, which has additional implications for possibilities for governance. Software studies and digital geographies have been productive in complicating understandings of human agency in relation to computation. On the one hand, agency is structured by how computation intervenes in processes of decision-making by representing places in particular ways (Mattern 2015). For example, predictive policing software (see Chapter 4 below) marks places as high risk for crime events at certain times, informing the deployment of police patrols. On the other hand, networked computers are able to extend human agency across space with unpredictable and often unseen effects. As Bratton (2015) observes, "[i]f a plastic stick in my hand can be used to control a video game character whose actions are calculated on a server inside a mountain drawing power from a nearby dam, then many even more bizarre chains of remote agency and distributed responsibility are possible" (338). Kitchin and Dodge (2011), following Adrian Mackenzie, call this "secondary agency"—a process that "supports or extends the agency of others such as programmers, individual users, corporations, and governments; it enables the desires and designs of absent actors for the benefit of other parties" (39). This extension of agency is an act of translation that is never perfect or direct, but it allows the encoding of agency into software to be enacted at another time (Introna 2011). It relies on the ability for people to control and influence computers that can then modulate and transduce space, implicating agency, software, and space in a complex and shifting network of relations.

Coded governance

In this dissertation, I use the term "coded governance" to describe how contemporary governance becomes imbued with the logics of software. Here, I understand governance as a distributed set of practices that are not necessarily undertaken by a central state or authority, but which come together to distribute resources, sort bodies, produce subjects, and make current social and productive relations possible. As I argue in Chapter 3, some practices that are initiated outside of official channels will be retrospectively enframed by the state, becoming embedded within official practices of governance. Computational modes of surveillance, for example, are sometimes developed and tested in ad hoc ways by hackers and tinkerers who attempt to see what is possible within current configurations of algorithmic infrastructures (see Chapter 2). These experiments with infrastructure—whether conducted within private software companies, academic labs, or on private computers—have important implications for political processes as they link to existing practices and are sometimes eventually adopted by state institutions. How software allows for governance to be practiced through unofficial and sometimes surprising channels, which later become unevenly integrated into official institutions, reveals important insights about how software is shifting political practices.

Extending earlier techniques of statistical techniques of governance, software affords the possibility to identify patterns across massive and disparate datasets, automate decision-making processes, and use predictive analytics to extend insights into the future (Amoore 2013). Recognizing the changing grounds of political processes, geographers have begun to identify how modes of "algorithmic governance" are changing how space is produced, understood, and governed as algorithms become embedded within infrastructures of governance (Crampton and Miller 2017; Amoore 2017). Written in code, algorithms can be tweaked and experimented with to produce far-reaching effects, constantly shifting how governance is coded. In the context of governance, coded is put forward here in three different senses.

First, coded refers simply to computer code, which enables computational and automated processes of governance. Once integrated into various political decision-making processes, the way code is written, used, and understood has important political, social, and economic consequences. For example, algorithms aimed at predicting gun violence have been shown to be ineffective, while leading to more arrests (Saunders, Hunt, and Hollywood 2016). Surveillance algorithms analyze social media data, interpret online browsing histories, scrutinize financial transactions, use video feeds to collect biometric data, automatically read license plates, and process audio, sometimes to add suspects to terrorist watch lists (Amoore and De Goede 2005), other times to target advertisements. And climate models, collecting data from a worldwide network of various sensing apparatuses, render climate change as a knowable threat to humanity on a global scale (Edwards 2010). The processes instantiated by code within these systems are often unstable and unpredictable as they combine with other processes in complex ways. Additionally, code is always being tweaked and updated, changing how software works across systems. Far from inevitable, each of the software systems above were produced through series of decisions, negotiations, and contingencies, which means that other models and directions are possible.

Second, coded refers to how computational processes of governance are often black boxed or otherwise rendered inscrutable. There has been growing concern over how public decisions are informed by opaque processes and assumptions with little or no regulatory oversight. As one software developer told me, “I don’t think we should be making policy decisions based on black boxes or opaque methods.” In response to coded processes, there have been calls for transparency and explanatory mechanisms to describe how insights are produced by computation.[5] But transparency runs up against the limits of private interests that want to maintain secrecy in the name of guarding trade secrets as well as the inscrutable nature of

[5] See, for example, the recently-formed AI Now Institute: https://ainowinstitute.org/. Also, New York City has recently passed a bill that creates a task force to study automated decision-making: http://legistar.council.nyc.gov/LegislationDetail.aspx?ID=3137815&GUID=437A6A6D-62E1-47E2-9C42-461253F9C6D0

popular machine learning techniques like neural networks. While regulation has been slow to catch up with the proliferation of software used in governance, there has been a growing movement of researchers attempting to reveal the hidden working of algorithms used to inform governance.

And third, coded refers to the tendency for software to classify people and places, sometimes using preexisting criteria, other times by producing new correlations and categorical groupings. Just as the Booth maps coded areas of London with their purported social classes, automated processes are frequently used to code people and places. Often influenced by existing epistemological orientations, algorithmic coding comes to reflect existing political priorities. The design of classification schema structures the production of knowledge (Bowker and Star 1999), often working against the possibilities for interpretive and situated modes of computing (Drucker 2009). In surveillance, that sometimes means coding a person as high risk for committing a terrorist act (Amoore 2011). In predictive policing, coding can mean marking a neighborhood as at high risk for experience a crime during a certain window of time. Coding the world solidifies ideological orientations towards space and people, making them knowable and governable not only in the present, but also extending into the future.

Constructing futures

With the proliferation of big data comes not only the possibility of revealing hidden patterns of correlation between various data, but also to extend those insights into the future. Machine learning, predictive analytics, and other computational methods allow for the future to be brought into the present and governed in the here and now (Amoore 2013). In some cases, these methods only sediment existing patterns and biases, making the future always a reflection of the past, while neglecting underlying causes that could work toward structural change. Real time data allows for predictions to modulate to changing conditions, making the future malleable within limited bounds. The futures that software can predict are always limited by the availability of data, as certain phenomena resist measurement and quantification.

Disparate modes of measuring, analyzing, and displaying data sometimes come together as a means of expanding the scope of what can be known and governed. Most notably in the "epistemological and methodological pastiche" of the urban dashboard of smart cities, where the city is defined and understood through visualizations of those things that can be measured (Mattern 2015). Instead of a seamlessly functioning and scientifically driven amalgamations of city data, in practice the smart city is more likely to be "assembled piecemeal, integrated awkwardly into existing configurations of urban governance and the built environment" (Shelton, Zook, and Wiig 2014, 3). The cobbled together sets of practices and systems, filled with contingencies and uncertainties, outlined in the chapters that follow are all the building blocks of the smart city imaginary. Not only imaginary, predictive policing and surveillance software is sold as part of urban dashboard packages, integrated with other visualizations of city data in prepackaged ensembles. How these slick visualizations representing the everyday processes of the city are built and how they deal with data raises important questions for how computation is shifting practices of governance.

Researching the methods through which data is produced is an important step in developing a deeper understanding of contemporary practices of governance. It is imperative to not only ask questions about data, but also about the algorithmic processes through which they are created, processed, and made knowable (Kwan 2016). Those who write algorithms, however, do not always know the provenance of the data they use, denying them the ability to ask critical questions about decisions, contingencies, and uncertainties that went into its collection. Users, on the other hand, tasked with using maps to inform policy decisions, often have little or no knowledge about how algorithms have produced the insights shown on screen. The futures brought into being through predictive analytics, then, often consist of discrete and non-overlapping methodological worlds that do not necessarily communicate between each other. This dissertation unravels some of the entangled socio-technical networks of computational systems to understand how data, algorithms, and visualizations are socially constructed in particular ways, which then go on to inform governance decisions.

Outline of the dissertation

Networks that coalesce around algorithms, making possible the integration of computation into modes of governance, are the focus of the three case studies that follow. I show how computation—acting through both official and unofficial channels—becomes integrated within existing infrastructures, surveillance apparatuses, and policing, which in turn, changes how each functions and is understood. Algorithms afford expanded abilities to extract patterns, identify correlations, track and surveil individuals, and reconfigure the spaces of everyday life. To see and be seen through the machine and its mediated visions of the world has rapidly become a nearly ubiquitous feature of daily life in major global cities.

In Chapter 2, I describe how existing infrastructure changes when it becomes infused with the control logics of algorithms. Drawing from lessons learned from other research projects as well as concepts from computer science, I describe three tendencies of algorithmic infrastructure: reconfigurability, interoperability, and deception/obfuscation. While computation is always built on an infrastructural base, algorithms change how existing infrastructure functions, making disparate systems interoperable. Computers make these interlinked systems controllable and modifiable, allowing for small code changes to have far-reaching impacts across multiple systems. Massive reconfigurations of algorithmic infrastructure can be implemented in ways that evade detection since algorithms regularly obfuscate their infrastructural base. This obfuscation is sometimes the result of intentional deception, but can just as easily be the product of the way that software is built. Rather than a readable, logical sequence of steps, software is usually built by piecing together disparate algorithms and software libraries, which can all be plugged into each other without ever having looked at their internal code. Software quickly exceeds any one person's ability to understand it, which presents challenges when it is used to control infrastructure and connect with other systems. Breakdowns, malfunctions, hacks, and deceptive logics all threaten the infrastructures that undergird everyday life. Algorithmic infrastructure, especially in its ability to link disparate systems, also opens possibilities for tracking and surveillance as sensors, databases, communication networks, social media, and other systems can be tapped and aggregated to produce new representations.

Disparate systems sometimes come together through creative and surprising means, revealing new possibilities for tracking and surveillance. Chapter 3 tells one such story, showing how an ad hoc online crowdsourcing effort to track the two suspects in the 2013 Boston bombing. After the bombing, users of reddit.com shared and analyzed photographs, cell phone videos, social media, news report, police scanners, and other data sources in an attempt to find the bombers. While the effort was ineffective, and sometimes disastrous as incorrect people were identified, users built complex understandings of the event space and revealed the extent to which disparate data sources can be aggregated to track and identify people. Inspired by the massive amounts of evidence that users submitted as part of the Boston investigation, the policing software and equipment company Axon—formally known as TASER International, makers of body cameras and electroshock weapons—has developed software tools to collect crowdsourced evidence. By producing formal methods to for police to aggregate user-submitted data, the software aims to integrate data collected from cell phones with existing databases, expanding the surveillance capabilities of existing digital sources during events. Data for this chapter was culled from archives of user forums where the investigation took place, showing the power of online discussions to affect an unfolding event as well as contribute to formal policing through ad hoc and informal practices.

Chapter 4 deals with more formal structures of policing, describing how predictive policing software is developed, understood, and deployed. Using a mixed method, qualitative approach, I describe how developers understand and relate to the software they build. Research began with six weeks of fieldwork in a private software company in the US that produces a number of geospatial software systems used for governance. During that time, I followed multiple teams working on different projects, attended planning sessions and meetings with clients, and met informally with employees in the hallways and during times between meetings. I met with and got to know a number of software developers working on various projects, observing how they collaborated, negotiated problems, and understood the potentials and limitations of the code they were writing. Through this research, I came to understand the culture and rhythms of this particular office and the daily practices that accompany the development and maintenance of their software projects. Chapter 4 emerged from my interest in a small team within the company developing and maintaining predictive policing software. Following early interviews and meetings with team members during fieldwork, I visited the office on two other occasions to conduct in-depth interviews with key software developers. These follow-up interviews were informed by further research I had conducted, which included reading promotional materials, technical documents, criminology, and policing literature in addition to attending a conference on evidence-based policing, going to two large trade fairs that sold policing technologies, and interviewing developers in the UK working on predictive policing software.

I begin Chapter 4 by describing how the turn towards proactive policing has facilitated the integration of geospatial software into police departments. Informed by evolving crime theories and enabled by the availability of various types of data, a number of predictive models have been developed that measure and represent crime differently. I outline how three prominent models function, attentive to the ideologies that guide each approach. I then outline four uncertain translations that are produced as these models are put into practice. They include translations between predictions and practice, insights and responsibility, data and predictions, and developers and the public. I argue that uncertainties that are endemic to predictive policing cultivate doubt amongst software developers, while opening important questions for critical scholars.

The dissertation concludes with summary of how these discrete chapters can form the grounds for future studies, leading to new questions that need to be asked about the role of software in political decision-making practices. To those ends, I suggest several creative and experimental approaches to studying software systems, which involve researching and intervening in existing code and data. I then use these suggested approaches to develop a proposal for future studies of predictive policing. The proposal that concludes this dissertation outlines future directions for my own work on policing, while aiming to inform others working in the growing field of digital geographies. It extends the goals of this dissertation as a whole: to expand the possibilities for understanding the algorithms that increasingly inform contemporary governance.

2. Algorithmic infrastructure

Introduction

> “we ‘primitive folks' worship source code as a magical entity—as a source of causality—when in truth the power lies elsewhere, most importantly, in social and machinic relations” (Chun 2011, 51).

The solid, sedimented, and tangled webs of cables, tubes, machines, and data centers give legible form to this strange thing we call “the digital”. A focus on infrastructure might give the appearance of a reprieve from the endless deferrals of meaning that each signifier floating across the screen brings and from the frustrations brought on by never quite knowing where to find the algorithms we write about. These material things give a form to the vaporous and ghostly functioning of digital media and algorithms (Chun 2011), a form amenable to representations in photographs and maps in attempts to chart out the extent that the digital landscape might be read. Infrastructure thus brings the messy and sometimes amorphous worlds of “the digital” down to the ground.

By giving a form and structure through which computation operates, infrastructure becomes a useful heuristic through which to understand contemporary geographies shaped by the digital. And as work in Science and Technology Studies (STS) insists, infrastructure is not limited to material things built by humans, but may come to include daily practices, norms, social organizations, standards, tacit agreements, ecosystems, bodies, and other things and processes that hold together a functioning system. Thinking of the digital or the algorithmic as infrastructure, then, forces us to produce rich accounts that attend to diverse things like human practices, cheap energy, capitalist social relations, and rare metal mining on which computation relies.

Seeing infrastructure as only a support for computation, however, misses a parallel movement as algorithms become integrated into the infrastructures that support them. This movement towards what I call "algorithmic infrastructure" and the theoretical and methodological challenges it poses is the concern of this chapter. Algorithmic infrastructure is produced as existing, non-digital infrastructure becomes integrated with algorithms as a way to control data flows and integrate these flows into existing systems. Algorithms, however, make existing infrastructures strange and difficult, inducing chips to perform millions of calculations per second at the speed of light, making and re-sorting voltage differences, connecting and disconnecting hardware, linking and delinking communications, making machines move and exert forces, inducing electricity to flow and pixels to flicker—processes moving so fast that we can only observe in limited glimpses. All infrastructure is complex and much of it is fragile, but algorithms can make infrastructure even more so, quickly exceeding the ability for its engineers to understand it. And so algorithmic infrastructure poses unique challenges for scholarship as something always evades being pinned down to simple taxonomic classifications. Attempts to classify the objects that make up computational infrastructure are quickly undermined by the complexities of the processes they support, which constantly resist representation (Mattern 2016).

In this chapter, I examine some of the ways that algorithms change existing infrastructure. Software studies, which I draw from to analyze this change, can be thought of as a practice of infrastructural inversion (Bowker and Star 1999, 34) as it brings forth the taken-for-granted computational processes that undergird everyday life. A focus on infrastructure expands the spatial imaginaries of software studies, forcing us to account for the complex spaces, relations, socialities, materialities, and structures produced by the digital. In what follows, I provide a brief definition of infrastructure as it applies to the digital. I then explicate three tendencies of algorithmic infrastructure that contribute to understanding the social and political implications of computing: reconfigurability, interoperability, and deception.[6] I conclude by showing how the interlinking of disparate systems produces individualized and differentiated experiences of infrastructures. The reconfigurability of these systems, however, reveals the possibility for other infrastructural configurations as tensions and resistance mark the politics of algorithmic systems. Taken together, these three tendencies of algorithmic infrastructure have important implications for governance and the construction of subjects as automated processes increasingly code people and places, producing automated modes of social and economic differentiation.

Defining infrastructure

> "The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it" (Weiser 1991, 94).

Infrastructure studies, observes Susan Leigh Star (1999), is "a call to study boring things"—the "hidden mechanisms" subtending daily life that are constructed through standards and technical specifications, organizational arrangements, and the built environment (377-379). While it is easy to take infrastructure for granted—as something commonsensical, obvious, or uninteresting—this would risk glossing over and missing the politics of infrastructure, as infrastructure regularly represents the "de facto standardization of a single, powerful group's agenda" (Star and Ruhleder 1996, 114). It also becomes easy to erase the vast amounts of labor that go into producing and maintaining infrastructure, work that is essential in allowing infrastructure to erase the means of its own production and disappear (Bowker et al. 2009, 9; Howe et al. 2016, 554). Always built on an existing base, studying infrastructure requires

[6] In this article, I draw from the computer science usage of the terms 'reconfigurable' and 'interoperable', but expand their scope to include the properties of infrastructures of computing. Deception also has resonances with the term 'obfuscation' in computer science.

attention to the labor that built it as well as the negotiations, challenges, and compromises that went into structuring it and continue to be expressed in its functioning.

Algorithms are made material as a way to move data, both existing as voltage or electro-magnetic differences in computer hardware.[7] In this chapter, I use the term "algorithm" as it is defined in computer science—an abstract description of a set of instructions used to process and transform data, often as a means to solve a problem (Wardrip-Fruin 2012, 17; Striphas 2015, 403; Gillespie 2014, 167). This means that these instructions are independent of their encoding in a particular programming language; instead, they indicate the processes that a computer will instantiate when code is run. For an algorithm to become infrastructural, then, it must be encoded in a language to become software, which can then be run to produce voltage differences in computational hardware, and later connected to an existing infrastructural base. Algorithms and data, then, cannot be divorced from their specific contextual and material instantiations, echoing foundational work in software studies (Hayles 1999) and more recent scholarship in geography (Kinsley 2014; Leszczynski 2014). Analyzing abstract descriptions of instructions also leaves open the possibility of making inferences about these often opaque infrastructures, signaling a necessary tension for the geographer of the digital who faces the limits of access and understanding. Studying infrastructure today, then, requires creative approaches to understanding geographies interwoven with invisible and shifting patterns of electricity that make themselves known through various computational interfaces.

Work in STS has identified a number of definitions and characteristics that can act as a starting point for understanding infrastructure—characteristics that will be expanded on later in this essay to address the unique properties of algorithmic infrastructure. In their influential work,

[7] Leading to Kittler's (1992) provocative claim that "there is no software."

Star and Ruhleder (1996) provide a series of concepts to help define infrastructure that include: “embeddedness”, “transparency”, “reach or scope”, “learned as part of membership”, “links with conventions of practice”, “embodiment of standards”, “built on an installed base”, and “becomes visible upon breakdown” (113, see also: Star 1999; Bowker and Star 1999). Consider, for example, recidivism software that is currently widely used in U.S. courts to guide sentencing decisions and, like a number of recent machine learning examples used to identify, assess, and categorize people (O’Neil 2016), has been found to reproduce racist decisions in sentencing (Angwin et al. 2016). Following Star and Ruhleder’s concepts: first, the software is *embedded* within an existing legal system with particular social organizations and practices. Second, it is *transparent* to use, supporting repeated sentencing tasks without difficulty. Third, its *reach or scope* extends past a single courtroom, making it implementable in multiple sites and states and bringing together diverse data sets. Fourth, it is *learned as part of membership* in the court system, usually by lawyers and judges, which also makes it surprising or unfamiliar to outsiders like defendants who can be surprised by their scores. Fifth, it *links with conventions of practice*, in this case, the practice of sentencing. Sixth, it represents the *embodiment of standards* as the software links to and interprets standardized criminal records. Seventh, it is *built on an installed base*, linking to existing court computers, communication networks, and criminal databases. Eighth, it *becomes visible upon breakdown* as defendants and lawyers notice the racial unevenness of the scores or when a court that depends heavily on the software cannot get it to work. And, finally, in an additional characteristic later added by Star (1999), it is *fixed in modular increments, not all at once or globally* as no one is really in charge (382). We see localized, modular reconfigurations in changing practices of sentencing in which judges interpret results differently; in differential practices of filling out the software’s questionnaire; or in tweaks to the algorithm in certain locations to reduce the racial bias of the system. These nine characteristics provide a useful framework for drawing out research questions for infrastructural software as an object of study.

Implicit in the fourth characteristic, *learned as part of membership*, is the notion that infrastructure is a fluid concept that exists in relation to situated practices, so one must "ask, when—not what—is an infrastructure" (Star and Ruhleder 1996, 112-113). That is because, "[o]ne person's infrastructure is another's topic, or difficulty" (Star 1999, 380). Infrastructure, then, is both an epistemological and relational concept that emerges in relation to practice and the situated positions of those who encounter it. So, for example, in the recidivism case, judges and lawyer might encounter software as a taken-for-granted infrastructure while defendants encounter it as an inexplicable difficulty that grants or denies their freedom based on unclear and confusing criteria.[8] Because "people make meanings based on their circumstances," they will encounter and understand infrastructure differently, leading Star (1999) to suggest an ethnographic approach to studying it (383). So too do developers have a different sense of algorithmic infrastructures as they tend to the creation and maintenance of code and hardware, including the constant uncertainty brought on by bugs and validation attempts, subject to ongoing negotiations with other developers, clients, governments, and stakeholders.

Many algorithmic infrastructures are held together and made possible through complex organizational arrangements and standardized practices (Bowker et al. 2009, 103). In the case of the internet, as with many other algorithmic infrastructures, its continued survival is surprisingly fragile, held together in part by a network of trust. The domain name system (DNS), for example, lies at the top of the internet's hierarchy, routing network traffic to the proper servers and holding the internet together. To keep this routing process consistent across the world requires a

[8] "I'm surprised [my risk score] is so low. I spent five years in state prison in Massachusetts," observes a defendant in the ProPublic story. (Angwin et al. 2016)

network of trust that assumes internet service providers (ISPs) and DNS servers are properly following protocols for domain name propagation (Galloway 2004, 10). When these protocols are not followed, whether through malice or accident, entire parts of the internet can be taken down as a result (Mathew and Cheshire 2012). This ability for algorithms to reconfigure the network on the fly is the first of three concepts that I introduce in this chapter, which will help explicate some of the unique properties of algorithmic infrastructure. Following Wendy Chun's (2008) suggestion that, "software can only be understood in media res—in the middle of things" (323), I begin with recent fights over internet access and control in Turkey and the agonistic politics of protocological controls that route network traffic to illustrate the reconfigurability of algorithmic infrastructures.

Reconfigurability

> "Acting as linguistic transducers, the coding chains impart astonishing power to even very small changes. Such amplification is possible because the constant reproduced through multiple coding layers is a pattern rather than a presence" (Hayles 1999, 31).

On 20 March 2014, Turkish Prime Minister Recep Tayyip Erdoğan vowed to "eradicate" Twitter in response to the website's role in facilitating anti-government protests and distributing wiretaps (Krajeski 2014). Hours later, faced with court orders to block the site, Turkish ISPs modified their DNS servers to make twitter.com inaccessible in the country. DNS is sometimes referred to as the phone book of the internet, translating Uniform Resource Locators (URLs) like "twitter.com" to Internet Protocol (IP) addresses like 199.16.156.198, enabling computers to establish networked connections. DNS servers, then, can act as chokepoints that grant or deny access to individual websites or the entire internet, hinting at the hierarchical or arboreal structure of a networked infrastructure that is far from decentralized or rhizomatic (Galloway 2004). Faced with modified DNS servers, some Turkish internet users found they could reconfigure their modems to use other DNS servers that correctly routed network traffic to Twitter's website. Indeed, one could purportedly find Google's DNS server addresses—8.8.8.8 and 8.8.4.4—spray painted on the walls in the Kadıköy district of Istanbul the following day.[9] Users' abilities to bypass the hijacked servers points to an important property of digital infrastructures: the ability for users and developers to quickly intervene in and reconfigure them. With a few keystrokes, users' network traffic shifted to a different network of material things we commonly associate with infrastructure—fiber optic cables, data centers, servers, and satellites—thus shifting the network topology of which they are a part.

The battle between the Turkish government and internet users was not over, as each tried to outmaneuver the other by using algorithmic means to reconfigure the network. The government, through local ISPs, was able to spoof Google's DNS server addresses, so users who used the 8.8.8.8 and 8.8.4.4 addresses were not actually routed through Google's servers, but through censored ones. Meanwhile, Twitter offered a workaround: users could send tweets via text message (SMS), thus bypassing the DNS block.[10] However, this would potentially expose activists to government surveillance within the cellular communication networks that were being used to bypass the DNS servers. Some users, by using virtual private networks (VPN) or the Tor browser, were able to route their network activity through other servers, making it appear that they were outside of the country, thus avoiding the DNS block and cell phone surveillance (York 2014). Algorithms became a central point of political agonism and contestation (Crawford 2016) for their ability to reconfigure infrastructure for political ends (Steyerl 2014).

The reconfigurability built into these networks affords the ability for users to use snippets of code, software, and hardware to make rapid reconfigurations to the network topology of

[9] https://web.archive.org/web/20140322000353/https://twitter.com/kaansezyum/status/446940008843206657
[10] https://web.archive.org/web/20140321051849/https://twitter.com/policy/status/446775722120458241

which their device becomes a part. So too does it allow for software developers to push code updates to millions of devices at once,[11] significantly changing how devices communicate with each other. Single coders, then, can precipitate enormous and immediate impacts on critical infrastructure. For example, by 2014 "two thirds of the Web relied on encryption software maintained by just one full-time employee," which was found to have a critical vulnerability (Eghbal 2016, 12). The potential impacts of code updates to an algorithmic infrastructure are sometimes made visible through the proliferation of bugs resulting from minor coding errors. For example, a single line of misplaced code[12] in an Apple security update broke SSL encryption verification for all iOS devices and many OS X users, opening the possibility for an adversary[13] to pass a bad security certificate and steal data from a user—a method that is believed to been deployed in Iran using another network vulnerability (Arthur 2011). Similarly, in an example given by Ted Striphas (2015), by merely changing a single database attribute for "adult" from "false" to "true", a technician at Amazon effectively delisted 57,000 books with a few keystrokes (396). Small textual changes in high-level programming languages can thus have enormous and far-reaching impacts on the world, causing algorithmic infrastructures to redirect data flows,

[11] In many cases, users must accept these updates, although some devices have settings to automatically accept updates.

[12] A snippet of the code in question is reproduced below. Each if statement has a "goto fail" command that is executed if a security check fails. The fifth line of code, however, is not encapsulated within an if statement, which means it is always executed, making the program bypass all later security checks, exiting this portion of the program without errors. See https://www.imperialviolet.org/2014/02/22/applebug.html for more details:

```
if ((err = SSLHashSHA1.update(&hashCtx, &serverRandom)) != 0)
        goto fail;
if ((err = SSLHashSHA1.update(&hashCtx, &signedParams)) != 0)
        goto fail;
        goto fail;
if ((err = SSLHashSHA1.final(&hashCtx, &hashOut)) != 0)
        goto fail;
```

[13] Adversary is used in computer science to mean an imagined actor who attempts to break into a system. The adversary is often used to illustrate potential security flaws in a computational system.

change network topologies, redistribute the visibility and accessibility of items in databases, and change the computational functions of millions of devices.

Previous infrastructure was also built to be reconfigurable—take railroad switches and telegraph switching boards, for example—but the complexity of computers exceeds our ability to predict what is possible in advance (Somers 2017). Often these possibilities are revealed in practice or in the moment of an event (Lally 2017) rather than forecasted through detailed risk assessments or engineering drawings deployed for other infrastructures. Part of this unpredictably can be attributed to the vast complexity of computational functions, while part is due to the possibility for creative code interventions by individuals to have far-reaching impacts.

Computer programmers and hackers[14] often harness the creative potential of reconfigurability to build new ways to imagine computational systems. The Tor browser's novel method of anonymizing network traffic by sending it through participating intermediaries, for example, reconfigures how a network functions in an effort to minimize the ability for users to be tracked using standard IP/TCP methods. So too do mesh networks, darknets, and other overlay networks take advantage of the reconfigurability of algorithmic infrastructure, producing other possible ways for the network to function, only allowing access to those who adopt the new communications protocols. These efforts produce their own algorithmic infrastructure that parallels, overlays, feeds off, or otherwise lies within existing infrastructural configurations. Reconfigurability, then, opens algorithmic infrastructure to other possibilities that detourn the computational, economic, and/or social logics of existing systems, making other computational functions and socio-technical networks possible.

[14] Here, I use 'hacker' to mean someone who experiments with computational systems to build new things. See, for example: http://catb.org/~esr/faqs/hacker-howto.html

Interoperability

> "No objects, spaces, or bodies are sacred in themselves; any component can be interfaced with any other if the proper standard, the proper code, can be constructed for processing signals in a common language." (Haraway 1991, 163).

If reconfigurability reveals the latent possibilities of algorithmic infrastructure, there is a simultaneous tendency towards interoperability as more digital objects use standardized communication protocols to connect to each other. The extent of this tendency and its global implications have come into focus recently as internet connected devices, often called the Internet of Things (IoT), have been deployed for massive attacks against internet servers. In one of the biggest attacks to date, video cameras, DVRs, routers, printers, and other networked devices[15] recently launched a massive distributed denial of service attack (DDOS) against Dyn, a DNS provider (Woolf 2016). The attack took down dozens of high profile websites for millions of users across the world (Neman 2016). These devices had been infected by the Mirai virus, a piece of software that scans the internet for IoT devices. It is able to infect, and effectively control, those devices with unsecured or default passwords,[16] which allowed it to launch the massive attack. It took advantage of the fact that algorithmic infrastructure has a tendency towards interoperability, allowing devices to connect to existing infrastructure.

Interoperability, in the case of the Mirai virus and in IoT devices more generally, is made possible by well-defined IP/TCP communication protocols that make the internet possible. This means that instead of a new infrastructure having to produce its own gateways to interface with existing infrastructure (Edwards et al. 2009), digital devices are able to "plug in" to the network using standardized techniques and methods. By using those same protocols, manufacturers can

[15] For possible list of devices, see: https://krebsonsecurity.com/2016/10/who-makes-the-iot-things-under-attack/

[16] https://www.symantec.com/connect/blogs/mirai-what-you-need-know-about-botnet-behind-recent-major-ddos-attacks

make it a simple operation, for example, to connect a home security camera to the internet and allow it to be accessed and controlled by a smart phone. Making those protocological connections is simple, while securing them is more complicated and time-intensive. In the interest of quick profit, then, a number of IoT devices take advantage of this easy interoperability, often at the expense of security. A piece of software like Mirai reveals the extent of the problem, as the IoT expands to become a vast algorithmic infrastructure of unsecured devices, vulnerable to novel methods of attack. The expansion of interoperable devices has opened the possibility for catastrophic failure of internet infrastructure (Wei 2017), leading groups like the Federal Trade Commission (FTC) to solicit security solutions to secure against the threat of attacks ("FTC Announces" 2017).

These attacks and their solutions also rely on the reconfigurability of algorithmic infrastructure as adversaries attempt to reconfigure these devices in order to create a flood of network traffic directed at specific targets. Software developers are tasked with reconfiguring communication protocols in order to secure devices against infection by viruses like Mirai. Both sides, then, address the global implications of interoperability by harnessing the power of reconfigurability, hinting at the spatial implications of algorithmic infrastructure as human actions, in the form of coding, have far-reaching and unpredictable consequences that extend across the globe.

So too does interoperability fuel imaginaries of seamless flows of information that traverse the globe. But, as anyone who has taken an introductory GIS course and learned to join disparate data tables knows, making these flows interoperable requires a lot of labor in the form of data culling, cleaning, and sorting. For advertisers and data brokers buying and selling users' browsing data, joining disparate datasets is a key component to constructing profiles of users and inferring their characteristics. So too do academics and cartographers, news media and city governments rely on making data interoperable, joining datasets to construct new understandings of the world. Some of these efforts are locked away in black-boxed, proprietary algorithms, like those used by data brokers. But other modes become part of the fabric of the internet, open to public use and experimentation. Application Programming Interfaces (APIs), for example, allow computer programmers to query a website and receive requested data in a standardized data format. Twitter, Facebook, and many social media companies have public APIs that allow this type of interaction, allowing programmers to create new algorithms that automate connections between social media data and other data sources.

The usefulness of data is often a function of its ability to be assembled together with other data, producing a more complex picture of phenomena. The table join is the precursor to the development of algorithmic means to automatically join data, enabling steady flows of seamlessly-integrated real time information. Or at least, that's how the imaginary of computing is narrated, and what better image is there than the smart city? A dream of giant screens in control centers, aggregating and analyzing disparate flows of data to produce a comprehensive view of the city—a city that can be efficiently controlled and managed.[17] But interoperability is produced unevenly and through the labor of coding and data management, a precarious achievement prone to vulnerabilities and requiring constant renegotiation as data protocols change and new databases are pulled into the fray, requiring the requisite flattening and integration into existing interfaces. Rather than a seamless world of data flows, we find people and machines working in concert, connecting and disconnecting flows, producing protocols for communication and

[17] "Interoperability doesn't happen by chance. As we deploy more and more sensors across city systems agreeing the basic standards becomes a critical component of its success. As we increasingly integrate systems we are also bringing different sectors together: the lamppost manufacturer, with the automotive company, the energy provider, the electric vehicle charging manufacturer, the (safety) camera technology provider, and data scientist. So working with the many different players in these new service value webs, to agree how best to design solutions in a more common way is vital if we are to open the market." (from: http://smartcitiescouncil.com/article/dont-leave-interoperability-chance-it-will-cost-you-big-time)

interoperability, tweaking thresholds and extracting patterns from the world to make it legible and governable (Amoore 2016), reflecting the infrastructural arrangements that make the "accidental megastructure" of computing hold together (Bratton 2015).

Flattened to the logic of computation and syntax of mathematics, everything becomes potentially interoperable. Algorithms used to interpret seismic waves probing the depths of the earth in the search for oil became auto-tune when aimed at reshaping pop vocal waveforms. Algorithms used to predict earthquake aftershocks became predictive policing when used to model patterns of "crime" (Mohler et al. 2011). Mathematical models can easily be applied to data, while well-known machine learning algorithms can be used to infer previously unknown correlations. By making everything commensurable, the world is remade in the image of the logics and tools we use to understand it (Hayles 2005; Rose-Redwood 2012). The tools that we use to measure the world can be aimed at nearly anything reduced to the binary logics of computation, giving machine learning algorithms, for example, a wide range of interpretive flexibility (Pinch and Bijker 1987, 40) that often exceeds anything imagined by their creators. If interoperability makes this world possible, reconfigurability undermines its inevitability, although how algorithmic processes work is routinely obfuscated, often intentionally.

Deception/obfuscation

While many infrastructures hide the conditions of their production and functioning as they become the taken-for-granted grounds of everyday life, algorithms are able to actively deploy strategies of deception to obfuscate[18] their infrastructural ground. By exploiting the gap

[18] This is in contrast to abstraction as a computer programming technique that hides some of the complexity of computational systems to make, for example, high level programming languages and collaborative programming possible.

of legibility dividing user interfaces from the actual functions of algorithmic infrastructure, algorithms are able to obfuscate their internal functions as a means to evade scrutiny, regulation, or accountability. In September of 2015, for example, the Environmental Protection Agency (EPA) revealed that the car company Volkswagen had been using sophisticated software to deceive environmental regulators. The software underlying the cars in question were able to use data about engine functions, speed, air pressure and steering wheel positions to determine when a car was being tested and reduce its emissions to meet regulatory standards (Hotten 2017). When not undergoing testing, the car could resume its normal functioning, emitting pollution at rates up to 40 times above emission standards ("Learn About Volkswagen Violations" 2015). Instead of these machines embodying standards as a means to integrate into existing infrastructure, Volkswagen was able to make them *appear* to meet standards through algorithmic functions deeply embedded within the computational and machinic functions of the vehicle.

The opaque and complex relations between digital interfaces, algorithmic infrastructures, and users can be understood, in many instances, as an intentional strategy of deception, often for profit. Just as Karl Marx aimed to uncover relations of exploitation behind systems of commodity exchange that were obscured by classical economists' misleading claims about capital—which we might read as a theory of infrastructure (Peters 2015,34)—uncovering algorithmic infrastructures often points to hidden systems of value production obscured by technological discourses.[19] For example, focusing an a user's role in generating online content would miss their more important role as a generator of profitable data (Van Dijck 2009). While connections between online participation and data accumulation for advertising and profit are well-known, this relationship is, of course, actively obscured by social media companies. The

[19] The connection between Marx's revealing of the hidden relations behind capital and the hidden material relations behind source code is the connection being made by Wendy Chun in this article's epigraph.

extent and functioning of these algorithmic infrastructures for profit is both constantly evolving and being revealed by researchers (Bujlow et al. 2015; Christl 2017; Thatcher, O'Sullivan, and Mahmoudi 2016)

Operating in a legal gray zone, the so-called ridesharing company Uber has also been found to use algorithms to deceive regulators. Through an automated analysis of user data—including proximity to government offices, credit card numbers, types of phones used, and other data points—Uber's smartphone software would display a fake version of the app to users it had tagged as potential regulators (Isaac 2017). These regulators, often tasked with enforcing a city's legal requirements for transportation services, would find it impossible to hire a car. Through deceptive practices made possible through algorithmic means, the algorithmic infrastructure of software-based ridesharing enforced a buffer between it and the infrastructure of regulatory apparatuses to which other transportation services are tightly integrated.

The use of deceptive practices in the constant search for profit, often paired with catastrophic social and ecological consequences, is a story as old as capitalism itself. Algorithms are merely the newest mode of exploitation and capitalist accumulation, leading to a shifting terrain of economic relations. Deceptive algorithmic processes are difficult to uncover, in part because code is considered a trade secret, hidden away in black-boxes, only discoverable through the analysis of inputs and outputs, through analogy with other systems, or through the occasional insider leak of code or methods (Kitchin 2017). Unlike many other infrastructures, algorithmic infrastructures cannot be easily cracked open or tapped to understand how it functions. In the case of Volkswagen, the deceptive software was revealed by a small university research lab doing exploratory testing of car emissions on the road (Glinton 2015).

While the Volkswagen example shows how powerful companies can deploy deceptive algorithmic practices on massive scales, there are numerous smaller examples of algorithmic deception that rely on the tendency for algorithmic infrastructure to be interoperable and reconfigurable. The Mirai virus's use of IoT devices for a DDOS attack and the use of VPNs to mask users' locations, for example, exhibit two possibilities for deception. So too does the rise of cryptojacking—a practice that relies on hidden scripts on websites that mine crypto currencies using the computational power of website visitors' computers (Goodin 2017), essentially converting visitors' home energy into money. And at the same time that Uber uses deceptive methods to trick regulators, Uber drivers have reported complex schemes to trick Uber's algorithms into inflating rider fares (Adegoke 2017) and paying drivers for faked rides (Cox 2017), siphoning off small amounts of money from the bottom through deceptive algorithmic means. One might read these techniques as counter tactics against Uber's use of software to obfuscate labor relations and avoid legal requirements for traditional employer-employee relations.

While algorithmic infrastructures are infused with the agonistic politics of attacks, obfuscation, and profit that deception makes possible, so too are official processes of governance susceptible to the very same logics. As political institutions increasingly contract with private software firms and use proprietary software packages, the same logics of algorithmic infrastructure as described above become incorporated into infrastructures of governance. Calls for transparency are overridden by corporations like Palantir and Northpointe who insist on guarding their trade secrets, while their software informs political decisions with little accountability to traditional legal structures. This has led to a number of efforts to research and reveal the inner workings of these increasingly powerful algorithmic infrastructures,[20] which only become known through limited glimpses that reveal their contours.

[20] For example, the recently-formed AI Now Institute whose aim is "understanding the social implications of artificial intelligence" (https://ainowinstitute.org/) and Privacy International (https://www.privacyinternational.org/), which investigates how governments and corporations gather and exploit personal data.

Consolidation and differentiation

At the turn of the 21st century, there was widespread fear that a programming shortcut would result in the breakdown of computational systems that had come to undergird so many aspects of everyday life. The Y2K bug, as it was called, resulted from programmers dropping the first two digits of the year, so "1999" become "'99" when written in code. How computers would deal with the turn to "00" was a cause for concern, as the turn of the millennium was an untested scenario—a known bug that had always been deferred to the future, but which never had been adequately addressed. Reflecting on the potential disaster that loomed on the horizon, Paul Edwards (1998) described the ways that computation had made possible the integration of so many heterogeneous systems. Previously discrete machines, Edwards observed, had come under the control of computers, making possible "integrated infrastructures of huge scale and scope" (16). With this integration, the failure of one system endangered others, as discrete systems like telephone systems, the internet, and electricity grids were all highly integrated and dependent on each other. If railroads systems were made interoperable by standardizing track sizes, explained Edwards, the ability for computers to translate between codes and protocols made it as if the tracks were infinitely flexible, able to integrate both existing and future systems into their logics.

The widespread integration of infrastructures continues to expand, bringing once disparate systems together. But computers do not only provide ways to control existing systems, they also enroll users in ways that were previously impossible. Each interaction with algorithmic infrastructures produces new data points in unknown databases that can be aggregated, analyzed, and used to personally target users. This individual targeting has paved the way for the introduction of neoliberal logics into an increasing number of services. For example, smart electricity grids, road pricing systems, biometric border crossings, and network data pricing schemes all work to commodify public systems in highly individualized ways (Graham 2005). Through these various processes of data collection and analysis, users are coded in databases based on inferred characteristics, which influences how they experience infrastructure in the future.

Computation thus allows for the extreme social and economic differentiation of people made possible by the consolidation and aggregation of diverse infrastructures. For big companies like Google and Facebook, the bundling of previously discrete functions allows for the massive aggregation of user data that can then be used to individualize and target services. Smaller companies, who do not have monopolistic control over communication networks, can use APIs and data brokers to access this user data, which can then provide the means to integrate services in novel ways. For example, the microfinance company Tala[21] offers small loans to people without credit scores in Kenya, the Philippines, Tanzania, and Mexico. To assess their credit risk, a smartphone application will access and analyze a user's text messages, phone calls, GPS location, contacts, and installed applications and make a loan decision based on the assessed probability of them repaying a loan (Privacy International 2017). Diverse datasets enabled by phone surveillance link social networks and everyday mobility patterns with financial markets, enabling the refinement of risk assessment. Interoperability offers the possibility of consolidating numerous data streams and molding infrastructures to algorithmically differentiate subjects. On one extreme are those subjects deemed risky or a threat. For example, some predictive policing software will target people based on their place within social networks (Saunders, Hunt, and Hollywood 2016). Or, as Simone Browne (2015) shows, automated categorization methods act to "reify boundaries, borders, and bodies along racial lines," building on a long history of racialized

[21] https://tala.co/

surveillance. On the other end of the spectrum are those deemed low risk and therefore not subject to barriers and prohibitions. The ability to bypass airport security, for example, enabled by background checks and biometric identification minimizes the friction of borders. Mobility, risk, surveillance, and governance are all differentiated and modulated through computational systems, influencing how everyday infrastructures are experienced and social relations are mediated.

Lauren Berlant (2016) defines infrastructure as the "movement or patterning of social form," which includes the structures, protocols, habits, and norms that bind us to the world (393). With the introduction of algorithms, infrastructures are able to code us and attempt to bind us each differently. But differentiation is always an act of comparison, coding subjects based on their similarity and relations to others. In the case of data-driven microfinance, probabilities are assessed based on comparisons with other historical risk profiles. In policing, risky subjects are marked as such based on those within their networks that are deemed similar. Processes used to code people and places produce difference rooted in networks of relations and measures of closeness, even if the specifics of those processes are often obfuscated. Following Berlant's suggestion that politics involves the reinvention of infrastructures that reproduce the unevenness of the current moment, here I argue that how difference is produced through computational infrastructures should be a central theme in that reinvention. Measures of similarity and difference and how they produce the subjects of the current moment can be the starting points for thinking about the reconfiguration of computational systems.

Examples of the reconfigurability of algorithmic infrastructures continually reminded us that other forms and possibilities exist. Latent within the all purpose calculating machine that is the computer comes other possibilities for standards, protocols, processes, and calculations that produce and modulate the conditions of everyday life. While current computation is built on top of existing infrastructures whose primary purpose is to accumulate capital, new protocols are sometimes built that resist that logic. Software that protects individual privacy, resists the commodification of everyday life, increases algorithmic transparency, or reduces the possibilities for mass surveillance might all be the starting points for reimagining computational systems.

Conclusion

The suggestion to consider algorithms as infrastructure is a call for the expansion of geographical imaginaries as a means to understand how the topologies of networked computation, operating through physical infrastructures of data centers and cables, are constantly reconfiguring material relations, extending human agency across space, and reshaping how we come to know and experience the world. Reconfigurability combined with interoperability hidden behind veils of deception results in the ability for machines to modulate difference in subtle and and often undetected ways. Algorithmic infrastructures increasingly produce the "ambient conditions of everyday life" (Larkin 2013, 336), shaping the ways we experience the world and each other.

Despite our best efforts to understand them, however, "there may be something, in the end, impenetrable about algorithms" (Gillespie 2014, 110). This is, in part, because of their tendency towards intentional obfuscation as explained earlier. But there are also forms of opacity that are produced through software design in various ways. Jenna Burrell (2016) outlines three type of opacity, which include: opacity as a corporate strategy to protect trade secrets, opacity resulting from technical illiteracy, and opacity resulting from the illegibility of computational processes. Burrell illustrates the last process through an image recognition task, providing a visual representation of how a computer "sees" text in order to classify it. These images appear random to the human viewer, but are meaningful to computer vision algorithms, illustrating the vast gulf between computer vision and human intelligibility. We would expect to encounter a similar gap of intelligibility between computational efforts to apply abductive reasoning to infer patterns in big data and human ability to discern patterns. Similarly, Wendy Chun (2011) argues that software is vapory and ghostlike, always resisting our best efforts to know it better. It is an important reminder to avoid falling into the trap of thinking that the mathematical logics that form the foundation of computing make it able to be known in some sort of absolute sense. Studying algorithms as infrastructure provide a similar lesson, as they can only be known through particular instantiations in code, integrated with other infrastructures, and embedded within a whole network of social and material relations. We can then expect any study of algorithms to be partial and incomplete, leaving them open to uncertainty, and doubt.

As we come to terms with the implications of living in a world filled with machines that produce and modulate space, it is important to continually confront the fact that they are always integrated into—and indeed, depend on— an existing infrastructural base. So while I applaud efforts to uncover the uneven results of algorithmic systems and work against those biases, we should not feign surprise when algorithms amplify the internal logics of already discriminatory and violent infrastructural systems. Rather, we should expect the introduction of computational means of enhancement to further solidify capitalist, racist, sexist, and heteronormative modes of sorting, accumulation, violence, and discrimination. The call for a theory of algorithmic infrastructure, then, is also a call for radical digital geographies that can both confront the existing social contradictions into which algorithms are integrated and use that knowledge as a starting point to understand how algorithms change those systems.

Accepted manuscript version, modified for this dissertation, reprinted with written permission from Sage:
Lally, Nick. 2017. “Crowdsourced Surveillance and Networked Data.” *Security Dialogue* 48 (1): 63–77. doi:10.1177/0967010616664459.

3. Crowdsourced surveillance and networked data

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). While the model is often applied to business projects (Howe 2009), it has also been used effectively in science (Bhardwaj 2014), policy (Brabham 201), 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 chapter, I focus on the /r/findbostonbombers subreddit—[22]a forum on reddit.com made for users to "compile, analyze,

[22] 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

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). 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).

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

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.

outcomes (Kitchin 2014), influencing how we perceive and think about the world (Gillespie 2014) "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."[23] 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).

[23] . http://imgur.com/a/sUrnA

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 contributes 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.[24] 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 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 judgments 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

[24] . 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.

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; 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). 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). 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.[25] 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) 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; Grusin 2010).

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

[25] 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

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). 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,[26] 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

[26] See https://rashomonproject.org/davis/ for a demo that uses the UC Davis pepper-spraying incident.

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), 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 phenomenon "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; Barocas and Selbst 2015). If this software is what structures relationships between people and the police, critical interventions become important.

We might also imagine myriad 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.

4. Translating predictive policing

Introduction

After fumbling with my smart phone for a few minutes and failing to produce a receipt of registration, the security guard leads me inside to a person behind a computer terminal. They greet me and ask for my last name, which results in a name tag emerging from a printer that is then slipped into a lanyard. I hang it around my neck as I am motioned towards the x-ray machines, looking down at the printed words that mark me as a “graduate student” from “The University of Wisconsin–Madison.” Despite rules stating that students are only allowed in tomorrow, and only those who have submitted proof of current registration in courses related to security (criteria that I would not be able to satisfy), no one seems to notice, or maybe care. Neither does the database that accepted my registration seem to have flagged the discrepancy, and so after retrieving my things from the x-ray belt, I walk into the exhibition hall. Over 300 booths fill the massive space, most of them private companies selling products to various types of security forces, from police and military forces to cybersecurity and border control agencies. I have come to see the many companies selling software that promises to assist and augment police and security forces. Shiny brochures, colorful vinyl banners, slick videos, and computational objects promise to make populations governable in a new ways and reveal hidden truths about the world.

The technology on display—which includes various types of computer vision, facial and object recognition software, drones and anti-drone devices, cybersecurity defenses, surveillance software, mobile technological command centers, and crowd modeling systems, to name a few—have all the makings of a dystopian science fiction future, but in this setting it is all rather mundane. In these halls, algorithms produced in the private offices of software companies come into contact with government representatives who shop for innovative approaches to policing. In the face of austerity, new threats, and the constant perceived need to revolutionize the instruments of policing, technology promises solutions in the form of more efficient distribution of resources and more precise targeting of populations. The trade fair offers a glimpse into how new technologies are talked about and sold, as well as how vendors attempt to push the latest gadgets and methods onto police forces. On display, the technologies in the show look impressive, making promises of smoothly functioning surveillance systems.

Technology promises scientific solutions to social problems, leveraging imaginaries of big data to sell a vision of a future that is knowable and governable. Automating surveillance and policing functions promises means to more effectively target and catch "bad guys" while minimizing individual bias and more effectively distribute limited resources. How these claims, and the data science that underpins them, match up to reality in practice remains open questions as limited studies have addressed specific technologies. Compounded by the fact that many computational systems are black-boxed—hidden away from view in the interest of maintaining competitive advantage for the companies that produce them—the unknown effects of software raise important questions for current shifts towards modes of algorithmic governance (Amoore 2017). But in these exhibition halls, how software functions and, perhaps more importantly, how it fails, is glossed over by adspeak as companies attempt to sell their products and enter into long-term contracts with government agencies.

The often secretive partnerships between public institutions and private software companies have begun to attract public attention. Recent articles in popular media, investigative journalism outlets, and academic presses have begun to question the assumptions and problems that come with using big data and software to inform governmental decisions (e.g., Lum and Isaac 2016; Kaste 2018; Ahmed 2018). Many popular articles focus on how using biased data to train software models will result in those biases becoming sedimented and replayed over and over (e.g., Robinson and Koepke 2016). Investigative reports have revealed how particular computational systems exhibit racial biases (Angwin et al. 2016), reflecting the racist social structures into which they have been integrated. Nowhere has this concerned been voiced more loudly than in relation to policing, especially in the context of growing public concern and awareness of the racist practices and outcomes of the police and justice systems.

Many critiques of software used to inform policing focus on data, tracing how data points are produced through everyday practices and transformed into actionable insights through algorithmic processes (Saunders et al. 2016; Ferguson 2017a; Degeling and Berendt 2017). Here I shift the focus to the sites of production and dissemination of software to see how developers stand in relation to these critiques.[27] By focusing on those who produce and market software—in the mundane sites of power of private software companies, academic offices, trade fairs, and policing conferences—I show how people, often removed from the visceral sites of power where bodies and forces come into contact, conceptualize, understand, and ultimately influence practices of policing. These developers, working together with machines, produce new ways of predicting the future, governing territory, and remaking the spaces of the city through novel

[27] This chapter is based on (1) formal interviews with seven people (which include academics and private software developers) who are involved with developing, selling, and integrating predictive policing software both in the US and UK; (2) site visits to a private software company where I sat in on several development meetings and talked with developers informally; (3) analysis of over 40 promotional videos that describe and sell predictive policing software; (4) analysis of academic, media, and promotional articles, many of them initially cited in interviews; and (5) visits to two security trade fairs and one policing conference. Since the number of people working on predictive policing is so small, all interview data is attributed to the anonymous moniker 'developer' and details about their workplaces and geographic locations are left intentionally vague to avoid re-identification.

Interview and video materials were analyzed to ascertain (1) how software functions, (2) what externals influences affect software production, (3) how developers view, understand, or imagine the world and policing, and (4) how problems are understood and negotiated. From this coding, the themes of this chapter emerged.

It should be noted that I do not claim that this study represents the positions of all developers working on predictive policing. Developers who will talk to an academic researcher are a self-selecting group and are probably more attuned to social and critical concerns than those companies who refuse any kind of scrutiny. They are also more likely to be active in public forums and partner with academics in the development of their software.

methods of data analysis. This study looks to shed light on some of these often opaque practices while describing how software developers and marketers understand and construct the world through the policing software they develop.

Instead of smoothly functioning systems like those promised in the trade show halls, I found cobbled together sets of practices and algorithms. Plagued by gaps of unintelligibility, unresolved tensions, hidden assumptions, and friction between actors, the networks that coalesce to build, test, and use policing software are just as fraught as any socio-technical system. It is in the gaps and voids that critical geographers can ask important questions about the political and spatial implications of software used to inform policing. Just as the database software of the high-tech policing trade fair mistakenly allowed me access to the technical wonders of automated computational analytics, there are numerous points of entry into the seemingly opaque systems that have come to increasingly govern everyday life. And just like the trade fair database, these systems are prone to errors and oversights that demand attention and open possibilities for critical research. As I walk through the halls of the trade show, I stop to watch a presentation on using social media data to monitor an emergency in progress. Someone walks up to me, and without saying a word, scans the bar code on my lanyard with a portable scanner, producing another data point in an unknown database. As data proliferates, so too do methods to sort, analyze, and make sense of it in novel ways.

In studying predictive policing, it is easy to dismiss these systems as ineffective—merely producing feedback loops that further entrench existing biases found in data. Instead, my approach examines the perspectives of those who create software and who have complex and evolving relationships to both the software they produce and the critiques leveraged at it. Part of this approach seeks to account for the effectiveness of software in understanding and governing the reality it has enframed, while simultaneously opening it to critique that does not dismiss the work and perspectives of the developers involved with its creation and use. The current climate of critique leveraged at software can easily fall into a similar cultural division that marked early work in Science and Technology Studies (STS). Evelyn Fox Keller (1995), for example, wrote about the seemingly unbridgeable communicative gap between scientists and the social scientists who study them. On one hand, she found radical critiques of science that did not account for the efficacy of science, on the other hand were scientists who relied on that efficacy without responding to critiques. Similarly, various software systems used in policing are verifiably effective at predicting crime—with all of the caveats that a concept like "crime" might entail, as explained below—which must be accounted for in developing critiques of these systems. Simple critiques are regularly confronted, addressed, and explained away by developers, crime analysts, and others who advocate for evidenced-based approaches to policing. In what follows, I analyze those counter-critiques to understand how developers of predictive policing understand and narrate their work and its effect. Throughout this chapter, I describe moments of uncertainty that emerge as problems with policing cultivate doubt amongst software developers.

Proactive policing and data analysis

The analysis of data has long been a part of policing, from the use of evidence in investigative research to the pinning of crime locations on printed maps to identify geographic patterns of crime. As computers and geographic information systems (GIS) became more accessible, computational methods to analyze and map crime entered police departments. Rolled out in 1995 in New York City, CompStat, for example, formalized methods of crime data storage, analysis, and mapping for police departments. Hotspot maps produced by the CompStat system identifying concentrations of crime would inform the deployment of police patrols in the city. A series of criminology studies, beginning in 1995, have consistently reiterated the effectiveness of combatting crime through the targeting of hotspots (Sherman and Weisburd 1995; Weisburd and Telep 2014), making it a widely adopted strategy in police departments (Committee on Proactive Policing: Effects on Crime, Communities, and Civil Liberties et al. 2018).

With the turn to computational methods comes the increased need for specialized technical positions within police departments. Many departments today rely on crime analysts trained in GIS to store, analyze, and map crime data and inform policing tactics and deployments. ArcGIS software, which is used and taught in geography departments around the world, has a range of applications and modules specifically tailored for crime analysis work, including hotspot mapping. Additionally, ArcGIS's parent company Esri advertises multiple events, tutorials, white papers, videos, and other promotional material related to crime analysis on their website.[28] The turn to hotspot mapping and other computational approaches to crime data was only the beginning of an ever increasing adoption of software by police departments, often supported by GIS methods and software.

The use of software to produce data-driven policing insights has been spurred on and supported by the turn to proactive policing. Facing a crisis marked by declining public trust, increasing crime rates, and the inability to prevent crimes, police departments in the 1990s began to move towards proactive methods of policing (Committee on Proactive Policing: Effects on Crime, Communities, and Civil Liberties et al. 2018). Instead of being a reactive force that only responded to emergencies and calls for service, this new method of policing called for strategic interventions that included hotspot mapping, community policing, broken windows policing, stop and frisks, focused deterrence, and problem-oriented policing. Whether focused on individual

[28] https://www.esri.com/en-us/industries/public-safety/segments/law-enforcement

offenders or geographic problem areas, these methods all rely on data to more precisely target people and places while generating data through police interactions, interviews, and arrests.

Proactive methods have become firmly entrenched within the day-to-day operations of police departments—new software packages promise to more efficiently and effectively hone these methods, often backed by claims of scientific methodology and objectivity (Ferguson 2017b). With policing once again facing a public crisis of legitimacy, most strikingly in relation to the high profile killing of a number of unarmed black men, but also in a number of cases of widespread corruption within police departments, this veneer of scientific objectivity can be perceived as a way to mitigate officer bias in policing decisions. Software simultaneously promises cost benefits for underfunded departments, making it easier for officers to more efficiently carry out police work and distribute resources (Perry et al. 2013). Efficiency here entails the precise targeting of places and populations, which often means being in the right place at the right time to effectively fight or deter crimes. Bolstered by academic partnerships with police to study crime patterns, federal government funding earmarked for technology acquisition, and constantly improving technology, computational methods were easily integrated with proactive policing (Ferguson 2017b). The move to a proactive ideology of policing, then, incentivizes data-driven targeting and preemption of crime.

If early data-driven approaches to crime using software analyzed existing geographic distributions of crime, more recent efforts have sought to predict or forecast where crime is most likely to happen in the future. Often called predictive policing, this mode of computational analysis uses crime theories and historic crime data to model crime in both time and space to produce dynamic and probabilistic understandings of crimes that have not yet occurred.

Geographic approaches to predictive policing,[29] which is the subject of this chapter, have important implications for how space is understood and governed by police forces. Often built by private companies in partnerships with public police forces, these software systems use a variety of methods and theories to produce models to predict crime.

On models, algorithms, and data

Crime has been forecasted using a number of different approaches, all underpinned by different data models and algorithms. While critical writing on predictive policing can often gloss over the details of how these systems work, it is worth being specific in describing the differences that each method produces. The decisions that go into producing each not only affects the technical functioning of the software, but also reflects the theoretical and ideological positions of its creators. The integration of data models and algorithms produce different possibilities for knowledge claims made about the world and have different implications for privacy, transparency, social justice, and the production of the spaces of the city. Each modeling process is the result of long processes of research and software development, leading developers to defend their models against others using various criteria that are built into the models themselves. For that reason, an understanding of how models work is paramount to understanding the intricacies of debates within the field of predictive policing as well how developers respond to critiques of their work.

Models of crime are produced through statistical or machine learning processes trained on historical data. Once produced, a model can then be used to analyze existing geographic and

[29] Predictive policing can also be used to target individuals to determine their likelihood of committing crimes in the future. For example, the Chicago Police Department produces a Strategic Subject List (SSL) of individuals determined to be at risk of committing future crimes by analyzing data on crime, social network, and other indicators with questionable results (Saunders, Hunt, and Hollywood 2016).

temporal data to output a probability of a particular type of crime happening in a given place and time. While hotspot maps are built on and represent existing crime data, predictive models are forward-looking, focused on emerging sites of concern rather than existing problems to be addressed. And while they are trained on data of past crime, they are generally dynamic, using constantly updating crime data to produce forecasts. Once a model is built, it can be tested on a selected number of recent months of data to test its accuracy in forecasting.

Methods for building geographical predictive crime models, including what algorithms and data they deploy, vary in complexity, legibility, theoretical foundations, and goals. As an emerging field of research, studies have not firmly established their efficacy as it translates to on-the-ground practices of policing, even if their predictive accuracy can be verified in relation to the data used to build the model. In other words, the data used to represent crime in these models exhibits identifiable patterns extending into the future, which can demonstrably be predicted with varying levels of accuracy through the use of models. But how predictive insights translate into practice and how data represents the phenomena it attempts to explain both remain open questions. Choices of models, algorithms, and data all influence what predictive policing systems are able to make visible in the world.

MODELS

Within the realm of geographic or place-based predictive police systems used within the US, there are a number of software systems currently in use in major cities (Robinson and Koepke 2016). The makers of place-based systems are careful to distinguish themselves from people-based systems, arguing that they focus on the *when* and *where* of crime without targeting individuals. This discursive move sidesteps some of the most glaring problems with predictive systems as they've been taken up in the media and popular imaginaries of technologically-augmented policing. The potential bias built into algorithmic decision-making leveled at individuals melts into space as geography becomes the risk surface to which these systems are trained, while time enables the precise targeting of crimes to come. In what follows, I outline three general types of place-based predictive policing models, all of which are in use in US cities today, sometimes in various combinations. They include near-repeat models, risk terrain models, and machine learning models.

Near-repeat models of predictive policing, which were the earliest attempts at predicting future crimes, are based on criminological theories of repeat victimization. These theories argue that places in the vicinity of certain crime types have an immediate increased risk of crime. These early studies, many of them originating in the UK, and the models that came out of them often focused on burglary data (Polvi et al. 1991). Arguing that burglaries are generally well-reported due to the need for police reports to fulfill insurance claims, as well as being relatively robust in terms of geographic, and to a lesser extent temporal accuracy, researchers recognized that existing data could be used to produce short-term forecasts of future burglary risk (Bowers et al. 2004). Near-repeat patterns that could be extended into the future promised to assist in the deployment of officers to at-risk locations. On the individual level, near-repeats in burglary are explained by similar housing containing similar and predictable levels of security. As burglars learn the layout, surveillance, and security of proximate houses, the argument goes, they feel more emboldened to target houses neighboring a successful target.

On an aggregate level, the phenomenon of near-repeat crime is explained in terms borrowed from public health and medicine—a way of understanding crime phenomena that structures how many predictive policing programmers think about and talk about crime patterns. Crime, in this conceptualization, is likened to an infection that displays a "contagion" effect as it is either transmitted to nearby locations or puts nearby locations at risk. To treat such a contagion, a “dosage” of officers is required to deter the future spreading of crimes. Dosage is often very precisely prescribed as a police patrol of 11 to 15 minutes at various times during the day, owing to a 1995 study by Christopher S. Koper that developed the “Koper Curve” (Koper 1995). The Koper Curve is frequently cited in predictive policing promotional materials, guiding how developers recommend using the insights produced through software. One developer, however, expressed concern with how entrenched this one study has become, questioning its validity to the wide range of practices and contexts that have become enrolled within predictive policing. In the movement towards evidence-based policing—which was inspired by evidence-based medicine (Sherman 1998)—attempts to ground practices and theories in scientifically-valid data is often undermined by the limited number of available studies. Those that exist often become the taken-for-granted grounds into which new practices are integrated. To create new studies and improve the science of policing, academics will partner with police departments to conduct randomized controlled trials (RCTs), which is the main method by which predictive policing systems of all types are validated. In nearly every interview I conducted with programmers, they would use the metaphorical language of public health and medicine to describe the patterns of crime they analyzed and predicted, as well as to describe how officers should deter future crimes. While these metaphors apply most cleanly to a near-repeat understanding of crime, these techniques and ways of describing crime weigh heavily on other place-based approaches to predictive policing.

While near-repeat patterns of crime were, and continue to be, mostly restricted to the analysis of burglaries and related crimes in the UK, the adoption of this theory in US-based predictive policing expanded to other crimes. Arguing that the contagion effect applies to gang violence in addition to burglaries, Mohler et al (2011) developed methods for predicting near-repeat patterns of these crimes. Using algorithms derived from earthquake aftershock modeling to forecast this contagion effect (Mohler 2015), the set of methods in this paper would develop into the commercial software PredPol. Using near-repeat patterns allows PredPol to boast that it only uses three data points to produce its models: "crime type, crime location, and crime date/time," which, it claims, eliminates privacy or civil rights concerns since no personal data is used.[30] By comparing emerging crimes with historical baselines, PredPol produces forecasts of future crime probabilities, which includes burglary, vehicle theft, robbery, assault, gang activity, gun violence, DUIs, and more.[31] Near-repeat patterns of crime, described through metaphors of disease, become modeled as a geophysical process that then shape governance decisions in a rather mixed set of analogies and metaphors built around theories of repeat victimization.

If near-repeat data focuses on the when and where of crime, Risk Terrain Modeling (RTM) combines similar data with environmental factors to produce an understanding of crime as it relates to these environmental factors. Premised on the idea that particular types of spaces and places increase the risk of crime, RTM claims to add the *why* to the *when* and *where* of near repeat models. Echoing the medical language of near-repeat, RTM claims to produce a "diagnosis of the environmental attractors of human behavior" by identifying correlations between chosen environmental factors and crime data.[32] The environmental factors that go into the model are chosen by the user and are meant to have theoretical significance to the crime (or other phenomena) being modeled. The outputs of an RTM model, in a sample case study on the RTM website[33], explains how the model might identify a correlation between gas stations and violent crime, prompting the police to patrol gas stations at particularly risky times. In the public

[30] see: http://www.predpol.com/how-predictive-policing-works/
[31] see: http://www.predpol.com/technology/
[32] see: http://www.riskterrainmodeling.com/about.html
[33] http://www.riskterrainmodeling.com/risk-reduction.html

health metaphors that abound, these risks are similar to the social determinants of health, here understood as the social and environmental determinants of crime.

Moving away from near-repeat models that are primarily concerned with time and proximity mapped onto an empty container of space, an RTM model inscribes a risk surface that is tied to chosen geographic factors. While connecting crime to geographic features might appear to bypass the social implications of predictive systems, it is important to remember that many geographic features that go into these models act as stand-ins or proxies for social or demographic characteristics (Degeling and Berendt 2017). Risk surfaces are built not only through individuated geographic variables, but also through the proximity of those variables to each other. Overlaying the map of a place, the risk surface that results classifies space as a variably risky surface based upon geographic attractors of crimes and their interactions. Adding time as a variable to these models results in a temporally shifting risk surface as the strength of environmental attractors might vary hourly or seasonally. The result is a dynamic space built through statistical measures of relationality focused on forecasting crime risk. In this model, risk is a product of the built environment, whether understood as attracting or affording particular behaviors.

Near-repeat and RTM models can be mixed in various ways to produce more complex models that attempt to describe crime (Garnier, Caplan, and Kennedy 2018). For example, some near-repeat models include street networks so distances along those networks, instead of straight lines distances emanating from crime event locations, are used to model crime diffusion as well as potential police interventions (Davies and Bishop 2013). Geographic features, like the locations of and types of business, are also sometimes added as an additional variable that can be used to find correlations with near repeat patterns.

The proliferation of big data combined with advances in machine learning have expanded the scope of claims to knowledge of crime in relation to forecasting as well as in its causal and correlative relations with other phenomena and spaces. Machine learning enables computational methods to recognize correlative patterns across disparate datasets and use those patterns to forecast future crimes. The abductive logic of pattern discovery on which these newer systems depend evades traditional hypothesis testing, instead tending towards an experimental and exploratory process that becomes difficult to falsify (Amoore 2016). In the machine learning paradigm, any data can be fed into the machine, whether it be the temporal and spatial data of near repeat patterns, the environmental variables of RTM, or any other type of data that might correlate with crime. A strict adherence to theories of crime becomes less important as the machine is able to analyze many possible correlations and determine which ones contain the strongest signals. IBM, Hitachi, and HunchLab all use machine learning in their predictive policing software while Motorola and LexisNexis likely do as well, although the latter two share very few details about their algorithms.[34] The shift towards machine learning approaches by many large companies poses a challenge to understanding how algorithms are implicated within political decision-making and governance.

ALGORITHMS AND DATA

It has become increasingly common to deploy the term "algorithm" as a shorthand to describe the functioning of a computer and it is worth explicating the place of algorithms in the models described above. An algorithm, as commonly defined in computer science, refers to a set of instructions used to transform data in order to solve a problem. Often likened to a recipe,

[34] see: https://teamupturn.gitbooks.io/predictive-policing/content/systems/

algorithms can be encoded in a number of different programming languages, which can then be compiled and run on a computer. In the models above, we can imagine how a dataset—points representing crime locations and times, for example—could be analyzed using a series of instructions to determine their spatial and temporal proximity, which could then be used to calculate a risk probability for that area, decaying over distance and time.

Near-repeat models of crime, which are based on a rather simple theory of criminal behavior, can be described algorithmically as a series of instructions to transform to data to match the theory on which they are based. Generally, models that are based on crime theories will have, at their root, intelligible algorithmic functions that can be broken down and described. This is because these types of models rely on programmers encoding theories into software, making them function in a way that is commensurate with the research that informs them. Machine learning, on the other hand, is implemented when there is not enough known about a particular behavior in order to model it. So instead of starting with a theory which can be modeled with an algorithm, machine learning algorithms extract patterns from masses of data, producing various types of templates of weighted parameters that describe trends in the analyzed data (Alpaydin 2016).

As programmers I interviewed explained, once a machine learning system is set up, the focus of the work shifts from the particularities of the algorithms that instantiate the machine learning analysis to the incorporation of data to be analyzed. This process is called "feature engineering," which is an essential part of adding new data to a model to make it interpretable by the machine. If the analysis of big data often starts without a theory or strong research question, instead looking for the machine to produce new insights about data (O'Sullivan 2018), it is through feature engineering that developers of machine learning approaches to predictive policing attempt to insert theories into their models. So, for example, programmers might add the locations of bars to their models to see if they produce a correlation with crime, theorizing that people gathering and consuming alcohol will lead to particular types of problems. Feature engineering is sometimes based on commonsense theories of crime like these, while other times it is based on specific research, usually coming from criminology. And sometimes features are produced or suggested by police departments adopting the software based on the intuition or experience of seasoned officers. In the machine learning paradigm, adding extraneous data that does not correlate to patterns of interest is usually not seen as a problem for programmers because the machine learning system will either disregard or give very little weight to data that do not correlate with crime. Because feature engineering is often a labor intensive process of making data interoperable with machine learning algorithms, there is incentive to include data that will significantly improve upon the model's predictive accuracy. There is, however, also some incentive for experimenting with different types of data that might reveal hidden correlations and processes, especially since there is little perceived harm in adding variables to the model.

As modeling shifts to a focus on feature engineering, it becomes important to understand what types of data are being fed into models. The rise of big data raises important ethical and epistemological questions as data changes how we can come to understand and govern the world (boyd and Crawford 2012). In the case of predictive policing, critiques often center around how models produce feedback loops as the use of policing data within models sediments existing structural inequalities. For example, using arrest data in models will mark already over policed areas as dangerous, which will lead to more arrests, further marking the area as dangerous. To avoid feedback loops like these, predictive policing vendors do not use arrest data within their models, despite many popular critiques that claim otherwise. Instead they use a combination of publicly-reported incident reports and calls for service, which are both subject to their own biases. For example, in the case of burglary data as mentioned above, those without insurance will be less likely to report crimes. And calls-for-service will likely be lower in communities that do not believe the police will help them. How crime data is deployed and understood, then, has important implications for whose interests are served by predictive policing software. Models will automatically determine correlations between crime data and various other variables in the system, which will necessarily reflect the unevenness of the original data. It remains an important question to ascertain those biases to understand how models are themselves difference engines that do important work in the world. Additionally, in some software systems, police departments are able to weigh the priority of particular crimes, thus producing uneven priorities that interact with the already uneven models of crimes, which are then integrated in various ways into uneven practices of policing.

Integrating into infrastructure

Technologies like predictive policing are unevenly integrated into already existing practices of policing and taken up differently based on existing structures and practices. Instead of making the world knowable and actionable in expected ways, predictive policing must integrate with existing social and technical networks of relations. These existing networks, as explained below, often evade scrutiny, complicating efforts to make software workable with existing practices. For the developers and marketers of policing technologies, this process of integration is a constant challenge that informs the production and communication of such technologies. As developers have explained to me, the initial trial period that precedes a long term contract is an important time for vendors to convince departments to adopt their tools. First rolled out in a limited geographic area of the city, trials allow developers to tweak their products based on department feedback as they strive to encourage officer buy-in[35]. There is a constant tension for the developers of policing software as they attempt to produce tools that can become easily integrated within departments while shifting how policing is practiced and understood.

Most predictive policing systems work by cutting a city up into grid squares, which vary in size according to the specific needs of a place. Future crime risk is forecasted for each of those grid squares[36]. The software will suggest the dispatch of officers to high risk grid squares— some software will even instruct them to engage with specific tactics, which might mean running patrol routes or talking with business owners—for a specified amount of time. Using software as a proactive method is reserved for times when officers are otherwise unengaged with other day-to-day activities, and is often directed by crime analysts within the department.

The proliferation of crime analysts within police departments following the adoption of proactive policing methods often acts as an entryway for the vendors of predictive policing software. Vendors will argue that their software can help analysts to sort through ever-increasing lakes of data and automate common processes in addition to improving accuracy. The move from hotspot mapping to more temporally and spatially specific forecasting is sold as a continuation of existing practices of analysis, not as a total reinvention. Crime analysts, in this narrative, are not replaced by automated computational processes, but are a necessary part of successfully implementing predictive systems (Perry et al. 2013).

To facilitate the integration of software into analysts' workflows, the technology is produced to be both simple to use and easily integrated into existing practices. To insure such interoperability, software is sometimes served from the cloud, accessible through either a web

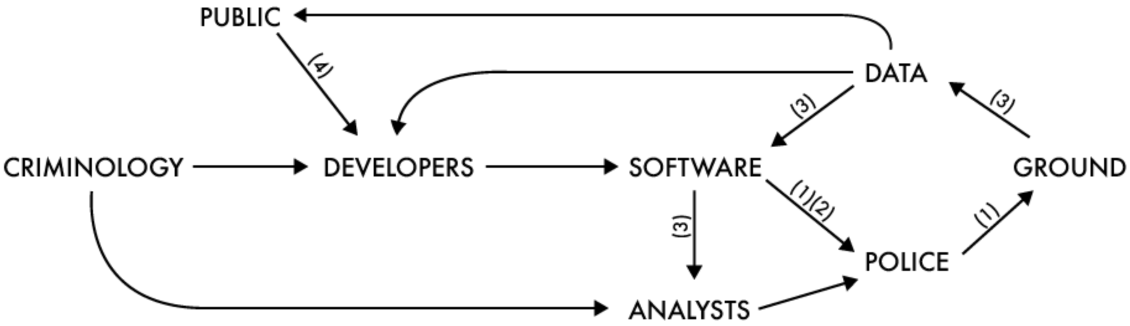
[35] Most trials are evaluated internally by the department, while some become evaluative RCTs. In Philadelphia, for example, HunchLab was tested in an RCT ("The Philadelphia Predictive Policing Experiment" 2018).
[36] See for example: http://www.predpol.com/how-predictive-policing-works/, https://www.hunchlab.com/features/, and https://www.motorolasolutions.com/en_us/products/command-center-software/records-and-evidence/commandcentral-analytics.html

browser or through an application programming interface (API). The former allows for the use by those without technical skills, while the API is able to serve data in formats that can be integrated into GIS software already in use within departments. The further analysis by analysts becomes important, as will be described below, when the machine reveals patterns without being able to provide explanations. It becomes necessary for analysts to do the additional legwork, acting as the human in the loop of computational processes, which both secures their positions within departments while deferring responsibility away from software companies.

Integrating technology into on-the-ground policing produces distinct challenges for integration. The dream of real time data streams augmenting police work and giving insights into constantly changing conditions meets real technical challenges in the form of equipment. Instead of tightly integrated technologies, officers are often given printed pdf maps at the beginning of their patrol shifts due to lack of available hardware. Cost barriers make it difficult for departments to purchase the hardware necessary to make predictive policing software integrate with practice as intended by developers. As one developer explained to me, software usually cannot be installed on in-car computers, since these are walled off from the internet, making it impossible for them to access online predictive policing systems. Instead, departments must buy smart phones or web-enabled tablets to access the software. It is an uneasy integration as cars are filled with more devices for predictive policing to function as expected. In interviews, developers have described a generational divide as younger officers are more willing to adapt to these changes, while older officers are more set in their ways and confident in their own intuitions to guide their policing decisions.

Four uncertain translations

The fraught and uneven processes by which predictive policing software becomes integrated into existing practices and infrastructures reveals a series of uncertain translations in which a number of important actors are rendered or choose to remain silent. Academic researchers and developers of predictive policing software, who are sometimes one in the same (e.g., Bowers et al. 2004; Mohler et al. 2015; Davies and Bishop 2013), produce prescriptive suggestions for these translations, arguing for how software *should* or *is* being used in practice. For example, based on a prediction, software might instruct officers to patrol a targeted area by foot for a set length of time. But these statements are often undercut by uncertainties that plague predictive policing as gaps in knowledge, or even what is knowable, limits what is able to be made legible. As Michel Callon (1986) observed in his analysis of actors within a socio-technical network, "to translate is also to express in one's own language what others say and want, why they act in the way they do and how they associate with each other: it is to establish oneself as a spokesman" (223). Through a series of transformations, various actors and data become enrolled within policing as a system, all the while being translated through a system of predictive policing. Software developers and academics, who are the focus of this chapter, narrate this translation, sometimes prescriptively, but other times descriptively. Despite their best efforts to sell the efficacy of their system and its potentially transformative effects on policing—especially in the case of private, for-profit companies—uncertainties in what is knowable haunt the implementation of such systems at every step of the way.



(Figure 1, translation identified through interviews)

In the sections that follow, I turn towards those uncertain translations, showing how they structure how software is narrated, produced, and understood. Uncertain translations are key to understanding the uneven implementation of predictive policing, but also reveal new entryways for asking questions about these technologies. If discussions between developers and critics often reach an impasse, as described below, it is the hope that critique can develop within these gaps, opening new questions that interrogate the current practices and potential futures that predictive policing produces. In what follows, I identify four uncertain translations that are inflected in software development practices (see Figure 1). They include translations between (1) predictions and policing practices on the ground, (2) ethical and legal responsibilities of programmers and police departments, (3) data used in predictions and their influence on predictions, and (4) public sentiment and developers of policing software.

Translation I: Between predictions and practices

The everyday practices of policing are notoriously difficult to study. As Mat Coleman (2016) argues, the "blue wall" that shields police practice is a formidable obstacle to studying the police, both qualitatively and quantitatively. He writes, "there is something perversely uneventful and chronically disappearing about police work which makes it exceptionally hard to excavate and interrogate" (77) as interviews become public relation events, observations do not disclose much, and power is forever disappearing. Quantitative measures of policing are generally unable to reveal bias in policing, while further abstracting from how power plays out on the ground (Coleman 2016; Woodward 2016). Despite hopeful and often imaginative claims for how predictive policing can transform policing practices, those who are involved with the production of such tools face a similar "blue wall" that limits their knowledge of how software is actually integrated into everyday practices. Resistance to changing practices within police departments is not a problem that only applies to technology, but plagues other efforts to introduce evidence-based practices to policing (Sherman 2015). Attempts at police reform meet the everyday and dispersed resistance of the blue wall.

Despite expressing frustrations when running up against this "blue wall," developers of predictive policing whom I have spoken to have admirable visions of how their software can precipitate positive changes to police practices. Developers claim that by increasing the accuracy of forecasts and sending officers to new places can contribute to reducing officer bias and over policing of communities. Instead of relying on infrequently changing hotspot maps that lead to constant patrols in the same neighborhoods, the dynamic, data-driven predictive policing maps, they argue, encourages more targeted and dispersed patrols that are ultimately more effective (Mohler et al. 2015). For some developers, the goal of shifting practices away from how they have always been done is a central to an ethical framework through which to understand predictive policing. If policing is facing a crisis, then experimenting with how patrols are directed and used makes sense as a reformist strategy. As one developer I talked to argued, you have to send police out somehow, so you can either use automated software systems or you can continue to do it in the ad hoc manner that it has always been done.

Within police departments, multiple challenges make software adoption difficult and attempts to integrate software within departments are often initially met with resistance by various actors. Overcoming this resistance and encouraging officer and departmental buy-in is a central problem in both the US and UK contexts. Officer resistance becomes a structuring problem that informs how developers create, test, and talk about predictive policing software. The efficacy of policing software relies on the often opaque practices of policing on the ground as much as it does on predictive accuracy. For software to show positive results, which is central to its long term and widespread adoption, effective policing practices have to coalesce around the use of software. For that reason, influencing everyday policing practices and encouraging integration within police departments becomes central to the software conceptualization, production, and testing processes.

While resistance is often measured through interviews with police and translated into software development strategies, occasionally resistance is experienced directly. Two developers, for example, described ride alongs they conducted with officers to observe how they were using the software. Both found the experience illuminating as it illustrated the concrete difficulties of integrating software into existing policing practices. They both described how experienced officers expressed skepticism about the software and during ride alongs showed no intention of using it. As one developer told me, an officer on a ride along made no effort to use the software, even though testing the software was the point of the ride along, instead he displayed his deep knowledge of his beat. "He's been there for twenty years and he just knows this thing like the back of his hand," described the developer, "and he almost… seems like a psychic when you're going around with him. He'll be like, 'all right I'm gonna go up this block, and then you see that person there, now if we go around on this other block he's gonna be there…roll down your windows a little, he's gonna say this thing as we pass.'" The predictive capacity of software, in this instance, is met with the predictive abilities of the seasoned officer whose intuitions and knowledge of his beat is used as a mode of resistance to data-driven technological augmentation. The sedimented practices of policing as it always has been done runs up against attempts to shift those practices, here exposed as a fleeting glimpse into how software might be summarily ignored even after being integrated into a department.

In another description of a ride along, a developer described how an officer ignored the predictive policing map because it usually just confirmed what he already knew, exposing a recurring problem that influences how software is built. Developers explain that software needs to display predictions that make sense to officers, which convinces them that it "works," while also revealing new things about the world, to show that the software adds value and helps reveal new insights about crime (Shapiro 2018). In this case, however, the officer discounted the software since it showed what he already knew, while he ignored those predictions that he did not understand. This led him to not feeling the need to patrol in the way prescribed by the software since it merely revealed what was already known[37]. Owing to the fact that policing is already notoriously difficult to study, these glimpses of resistance to software raise important questions for the possibility of software ever substantively changing policing practices in ways envisioned by software developers especially considering the current climate and culture of police departments.

To encourage officers to comply with new protocols initiated by software, developers will deploy strategies that (1) exploit the hierarchical structure of police departments or (2) attempt to account for individual psychology when building tools. In the first strategy, developers recognize the top-down power structure of policing as orders from commanding ranks are expected to be

[37] The officer, in this situation, did not perceive the software as being useful to his job, which is a major reason for the rejection of new technologies according to the Technology Acceptance Model (TAM) (Davis 1989). Some developers I interviewed use TAM as a way to increase buy-in from officers.

followed by lower ranking officers. By convincing top brass of the efficacy, fairness, or cost saving potential of predictive policing, the hope is that their buy-in will convince others to follow suit. So, for example, a chief of police in a major city might call for the adoption of predictive policing software, convincing sergeants to organize their precincts using the logic of that software. Within the structure of the department, belief in the software combined with the hierarchical command structure will presumably trickle down through the ranks. In this model, it is important for vendors to convince top ranking police of the value of software. Additionally, buy in from top brass can lead to crime analysts using software in day-to-day operation, which has an effect on how officers are distributed.

Even if software is integrated into the command structure of a department, it may still face resistance by individual officers who reject it on various grounds. A quick scan of online forums and blogs dedicated to policing reveals a plethora of suspicion and abhorrence directed at predictive policing, especially among those who claim to have used it. The problem of resistance leads to the second strategy for integration, which entails attention to individual officer psychology when building software tools. In interviews with US and UK developers as well as in literature, developers and researchers identify psychological reasons for officers to reject software, which includes skepticism, boredom, and lack of agency (Shapiro 2018). While software companies and academics will work closely with police departments while software is being integrated into the department—which includes soliciting feedback from officers through interviews, surveys, and ride alongs—how those efforts lead to changes in practices is always foreclosed, in part, by the always disappearing nature of policing. But officer psychology does play an important role in software production as developers attempt to maximize incentives for software adoption while addressing ways that software functions might block adoption. For software to be successful and appeal to officers, it is thought, it must conform to their own intuitions of crime while showing something new, allow agency of the officer in making decisions, and avoid officer boredom.

The first concern—showing something new while meeting expectations—is narrated as a central problem of predictive policing that affects how both officers and top brass come to trust a new system while seeing that it adds something important to policing (Shapiro 2018). In many ways, this is built into these systems, since a working data-driven approach should, in theory, accurately reflect the phenomenon it is trying to model while not being beholden to well known problem areas (Bowers et al. 2004). The second concern of allowing for officer agency can be addressed by giving officers choices of what to do if they are assigned to patrol a grid square (Perry et al. 2013). Instead of just patrolling the streets of that area, software and departments will give officers tactical choices, that might include community outreach, data gathering, or other active problem-solving approaches. The final concern of officer boredom might be the most difficult as the narration of how predictive policing is meant to work runs up against the idea of officers as active, crime-fighting agents, instead posing them as crime deterrence scarecrows, as one developer described to me.

In interviews with developers as well as in promotional material, predictive policing software is, rather counterintuitively, almost always narrated as a means to deter crime while lowering rates of arrests. Instead of "catching bad guys"—which is central to how officer desire is narrated—various software packages are meant to send officers to the right places at the right times and, due to their presence in a particularly risky place, will prevent crimes from ever happening. Under this model, policing becomes devoid of proactive strategies as crime merely disappears, leading to concerns with officer boredom and feeling like their work is not making a difference. Some of these concerns are mitigated through software by deploying officers to new places, suggesting novel tactics, and engaging officers in a problem-solving approaches as described above.

But how officers act on predictions and the effects those actions have on places that are deemed risky remains under studied. As one developer told me—which was echoed in other interviews—the unknown effects of officers on the ground, "is a big elephant in the room about all of this predictive stuff—about how well it really translates into results on the ground. We can improve our predictive algorithms all day and maybe get them pretty good, but whether a good prediction results in the reduction of crime and how to actually achieve that in terms of behavior on the ground is kind of an unknown at the moment." The opaqueness of policing practices compounds this problem as what happens on the ground and how it matches up with actions prescribed by software, analysts, or command centers is perpetually evading capture and subsequent analysis. For some software vendors, the possibility of tracking police to see how they interact with software could fill in some of these gaps, but it remains to be seen of such transparency will be palatable to officers and departments. Some departments have put up resistance to data collection methods intended to surveil or evaluate police practices (Kaste 2018).

In meetings I have observed and in interviews, developers will attempt to close this gap by imagining how officers might interact with new features. Groups of stereotypical users are used as devices to imagine how a new software feature might be used. Built in part though feedback from police, these groups might include officers who might be receptive to software, not interested, or absent-minded, with each establishing a different relationship to the machine, which in turn produces different outcomes and practices. In imagining users, developers attempt to identify with the everyday practices, concerns, proclivities, and feelings of officers in the field in an attempt to bridge the gap that separates the functioning of software with how users actually use it. When testing new features, developers will even get into their cars and pretend they are police, allowing the software to lead them to places of increased risk and patrol as prescribed. As one developer told me, when writing code, it can be difficult to know things will play out in practice, so using the software is a means to develop spatial awareness of particular use cases and, in turn, make software features more usable.

In idealized notions of predictive policing, officers will use software to be in the right place at the right time, thus deterring crimes from happening. Patrols in those areas will engage in community relations and data gathering, while building trust with communities. Software will reduce over policing of areas by precisely targeting police resources where they are needed the most. But against this idealized model used to sell predictive policing to both departments and the public, there is always an undercurrent that recognizes the negative impact of policing, which constantly underlies these discussions. Utopian visions of what policing could become are undermined by glimpses of actually existing practices on the ground, leading software developers to distance themselves from potential abuses and misuses of predictive systems.

Translation II: Between insights and responsibility

Idealized versions of policing run up against the racialized violence of policing in practice, raising questions about how software will either reinforce, undermine, or function alongside existing inequalities. Rather than being ignored or explained away, problems with policing are recognized by developers who deploy various strategies to address how software fits into this notoriously fraught field of practice. While critics argue that predictive policing will reinforce existing biases in policing, developers deploy careful strategies to minimize their complicity with oppressive practices. Strategies include deferring responsibility onto police departments for potentially oppressive uses of software, adding friction to software to limit how it might be misused, and prescribing normative practices detailing how technology should be used.

In an academic paper, which includes two of the founders of PredPol among its co-authors, researchers deployed an RCT to determine if PredPol leads to biased arrests (Brantingham, Valasik, and Mohler 2018). Recognizing that racial disparities in policing are rampant (citing 27 studies), the authors tested to see if racial biases were heightened in areas that software had marked as high risk. While overall arrests were not affected, arrests in areas that were marked by software increased significantly. Additionally, racial disparities in arrests remained unchanged as arrest rates rose across the board in predicted areas. The authors hypothesize that arrests increased due to predictive policing software being more effective at predicting crime than existing policing practices, citing their own academic study of the software they build and sell, which happens to be one of the few studies that has shown the effectiveness of predictive policing (Mohler et al. 2015). The findings are significant because they undermine the idea that predictive policing will lead to deterrence without increasing arrests rates. Instead, according to this study, software geographically targets predicted areas for increased arrests while maintaining existing arrest disparities.

How officers act upon forecasted areas of risk remains a matter of concern for developers and legal experts. As Andrew Ferguson (2017b) has argued, how predictive policing factors into reasonable suspicion or probable cause judgments that would justify a police stop remains legally unknown (see also: Perry et al. 2013). Citing the Supreme Court ruling in Illinois v. Wardlow (Rehnquist 2000), Ferguson shows how a person being in a "high crime area" was used as admissible evidence for police to determine reasonable suspicion and stop a suspect. In over policed communities, a person being in a "high crime" area is often used as a justification for unwarranted stops and searches, often of people of color (Kaufman 2016). Ferguson (2017b) argues that predictive policing creates mini high-crime areas in which "police may feel additional license to investigate more aggressively" (79). Whether or not this increase of scrutiny driven by predictive policing violates people's 4th Amendment rights against unreasonable searches and seizures remains legally untested and will depend on how Illinois v. Wardlow is interpreted in regards to crime predictions. In talking about software, vendors will reiterate that their forecasts do not translate to reasonable suspicion, and should not be used to justify stopping or detaining people in an area deemed to be high risk. This shifting of responsibility onto officers runs up against the realities of opaque policing practices, where the translation of predictions into action constantly evades capture. The gap between deterrence as it is narrated by developers and the possible reality of increased arrests and potential racial profiling within prediction areas raises important questions for how software might redistribute and concentrate the geographies of policing and arrests.

One developer told me that there will eventually be a court case, probably arising from a police stop that goes terribly wrong, where the role of software comes into question. How that scenario plays out has important implications for the future of predictive policing and its role within police departments. If predictive policing is about shifting how the world is seen by police, associating new spaces and the people within them as crime risks, how these predictions are understood has important constitutional implications for people who become the target of data-driven policing actions. Embedded within the name "predictive policing," argue developers, is a fundamental misunderstanding of the technology, which frames how officers might relate to it. Instead of producing predictions, vendors will often prefer the term "forecast" or "hunch" to describe the insights produced by their software because "prediction" is a concept that can easily be misunderstood. In fact, predictions are outputted as probabilities, which are very small. For example, as one developer described, probabilities of a crime happening within the small amount of time allocated for patrol in the highest risk areas is usually only as high as one or two percent. But the term "prediction" can be interpreted by an officer as a foregone conclusion of crimes to come, leading to overreactions against and profiling of people who happen to be in the boxes marked as risky. In addition to making statements prescribing how their software should be used, developers will program features in ways that reduce potential misuse or add friction to work against bad practices. For example, one company has recently removed statistical descriptions from prediction boxes, limiting what an officer in the field can know before patrolling an area. This change was in direct response to potential 4th Amendment violations. Developers reason that a box showing a high likelihood of a particular crime happening could lead to reactionary policing of people in high crime boxes. Instead, as the argument goes, the box should lead an officer to patrol a particular area and produce a deterrence effect. By tweaking what information is available, developers hope decision-making will be nudged towards the intended uses of predictive policing, even if the differences in map reading tasks have not been tested (cf. Vincent et al. 2018).

In another feature alteration, a police department requested manually determined patrols to be added to automated patrol allocation software in cases where officers had intelligence that would not show up in an automated system. This request, in the view of developers, opens their software to potential abuses as it increases the decision-making powers of officers who can then override predictions for arbitrary or nefarious reasons. Ultimately, developers succumbed to the request, but limited its functionality. In the implemented version, manual patrols cannot be recurring, can only be altered in the beginning of a shift, and they are clearly differentiated from automated patrols both in their map symbolization and in statistics kept in the system. For one developer I spoke to, negotiations to determine how features are implemented seemed to be an important part of maintaining a sense of agency in recognizing and limiting potential ethical problems with software as it is put into use. It is an illustration of the ambivalence of developers as they produce tools that they believe contributes to making the world a safer and fairer place, while recognizing the profound problems of policing and potential threats of a technologically augmented surveillance state.

While developers deploy small programming tactics to increase friction and produce ethical visions of how software should be used, the fluidity of software in its implementation opens it up to untold abuses. Making software adaptable to different policing contexts is a central concern for vendors who want to sell their software to police departments with different needs, practices, cultures, and workflows. To do so, software is produced to be reconfigurable in its implementation and interoperable with other software and data (see Chapter 2). So while developers will address ethical issues by prescribing what kinds of crime data is appropriate to model, how patrols should be conducted, and how software should integrate within departments, the fluidity of software means all of these normative suggestions can easily be undermined by those who use the software. For example, reconfigurability in software allows for users to incorporate their own data into models, data that may or not meet the ethical standards of either developers or the criminology literature they draw from. Adding unethical data, however, might fit into current departmental practices and appear to be effective when using current measures of policing.

In a study conducted on Oakland crime data, for example, Kristian Lum and William Isaac (2016) show how training PredPol's algorithm use drug arrest data would result in targeting black people at around twice the rate of whites, despite drug use across the two groups being relatively equal. Objecting to the premise of the study, Andrew Ferguson (2017b) points out that PredPol does not use drug or arrest data in producing its predictions because they, as do most software vendors, recognize the racial biases baked into that data. In response to Ferguson, Issac and Lum (2018) write that it is rather fanciful to think that, "predictive policing vendors will be able to self-regulate merely by voicing concerns about the potential harms of these tools." Citing government pressure to expand predictive policing into ethically questionable realms, the authors point to a National Institute of Justice "Real Time Crime Forecasting Challenge,"[38] which included drug crimes among other smaller, often officer-initiated crimes that are not used in current predictive policing software. For Isaac and Lum, the inclusion of the co-founder of PredPol and the chief data scientist of HunchLab signals the willingness of such companies to include questionable data within their models, even if it goes against current recommendations. Responding via twitter, HunchLab distanced itself from the contest, saying its data scientist participated in a personal capacity, also commenting that drug arrest data was not used, rather, all data came from calls-for-service[39].

Efforts to incorporate questionable data are perhaps most revealing not in showing how software vendors might be inclined to violate their own ethical frameworks, but rather, in illustrating how easy it is to misuse these systems. A recurring theme in talking to developers is how easy machine learning can be reconfigured and misused. As described earlier, machine learning algorithms are selected from a small range of choices and tweaked in various ways to improve efficiency or accuracy. Once set up, it is both easy to incorporate new data and act on predictions on ways not intended by developers. In fact, this type of reconfigurability is often built into the systems to make it attractive to potential clients. One developer, in discussing the ethical questions faced in developing predictive policing, warned of "how dangerous it can potentially be in the wrong hands," if, for example, "you do it the wrong way and come up with some Draconian vision." Even without misusing data, problems arise in how predictions are implemented and acted upon. One developer described how Immigration and Customs Enforcement (ICE) could use existing predictive policing software to target undocumented

[38] https://www.nij.gov/funding/Pages/fy16-crime-forecasting-challenge.aspx
[39] https://twitter.com/HunchLab/status/995825813508558849

immigrants who get robbed because they carry cash, switching the purpose of the software away from protecting people to targeting people.

Adding to concerns over misuse are the ways that software is casually talked about when being promoted or sold. For example, developers will openly talk about how the addresses of parolees and others recently released from prison could be incorporated into existing models. Or, conversely, models could guide recidivism decisions by analyzing the crime risk of a particular area to which a prisoner is released. The reconfigurability of software facilitates the addition of any data, whether or not it meets some ethical criteria, while its interoperability allows it to be easily integrated into other software systems through API calls. ShotSpotter, license plate readers, CCTV cameras, and other kinds of surveillance technologies can, and sometimes do, become integrated with place-based predictive policing systems, compounding potential issues with software. In the isolated conditions of the software laboratory, ethical problems can be summarily addressed and dismissed, but once integrated into the messy and amorphous world of policing practices, technological systems, and messy data, software opens up to a fraught field of intersecting oppressions and abuses that exceed the accounting abilities of software developers. Deferring responsibility becomes a key survival strategy for software vendors, but reaches its limits when met with the unresolved voids in discourses and practices.

Translation III: Between data and predictions

Many of the concerns with predictive policing discussed thus far have revolved, in one way or another, around questions of data. How well data reflects the reality it claims to represent, how data becomes actionable through predictive analyses conducted by machines, and how policing practices, when conducted in concert with machines, can change the data that goes into predictive analytics are all central questions that remain unresolved in many ways. For both software vendors and their critics, these questions lie at the heart of contested understandings of predictive policing and all structure how technologies are built and understood.

For police departments, the question of software's effectiveness guides decisions on whether to buy, and later continue paying for predictive policing systems. Departments demand statistics that show dramatic crime reductions that can justify their investments in software, both in terms of the money required but also the time investments in training and changing practices. The desire for numbers validating effectiveness in the short term often runs up against the longer time frames that are needed to produce in-depth validation studies. For some companies, producing validation measures within the software itself serves as a means to begin addressing the short-term needs of departments. But RCTs still remain the gold standard for validating software used in policing, even if they usually require partnerships with academics as well as time frames measured in years instead of months.

Attempts to validate predictive policing will focus on two facets of the software: its ability to accurately forecast crime into the future and its effects on those crime rates. The former is relatively straightforward and involves training a mode on several (often five) years of historical data and seeing how well it matches up with the most recent three months of data. This accuracy validation is often used to sell predictive policing, indicating the geographic and temporal precision that departments can use to focus policing resources. The latter, however, is more difficult to measure, but becomes more important after a department has implemented software for a given amount of time. For Aaron Shapiro (2018), how predictive policing leads to interventions that change the crime dynamics it attempts to represent becomes a structuring problem for developers. Predictive policing, as described above, is premised on a model of deterrence, which, as the argument goes, means patrolling in areas deemed to be risky deters crime and, as a result, destroys the patterns of data that went into making the risk prediction.

Proving the absence of crime is statistically difficult, while confounding factors in the real life scenarios in which RCTs are implemented further complicates the ability to validate software (Shapiro 2017). Additionally, if crime is deterred in one place, it could diffuse into other adjacent areas.[40] In fact, only one academic study has shown statistically significant drops in crime associated with the use of predictive policing—a study that was conducted by the co-founders of PredPol and that used their software (Mohler et al. 2015). Other studies, including one by RAND, have found no statistical differences coming from the use of predictive software (Hunt, Saunders, and Hollywood 2014). It is perhaps too early to draw conclusions about predictive policing's effectiveness, and even if studies to show a correlation with crime reduction as it has been defined, there are other fundamental questions that can be asked about data.

Data that goes into models always poses a guiding problem in predictive policing as it is widely recognized that it can never provide an accurate picture of crime as it exists. The task, then, is often to minimize biases that would lead to uneven or oppressive impacts. For example, developers widely recognize that drug crimes are largely officer-initiated, reflecting patterns of over policing and racial bias more than actually existing patterns of drug use. Major software vendors, for that reason, do not use drug arrest data or any other officer-initiated data like arrest data. Using arrest data would lead not only to the sedimentation of existing patterns of bias, but it would also result in feedback loops in which officers arrest people, that arrest data goes back into models, which then leads officers to visit the same places and arresting more people. To minimize the possibility of both bias and feedback loops, developers will use a mixture of incident reports initiated by the public and calls for service.

[40] Despite difficulties involved with measuring diffusion effects, there is some evidence that there is a diffusion of crime deterrence benefits that arise from hotspot mapping (Weisburd et al. 2006). No studies that I know of have shown statistically significant results showing crime diffusion effects of predictive policing.

Both calls for service and incident data are known to be biased, but developers I have spoken to argue that those biases reduce existing patterns of uneven policing. For example, on aggregate, minority and immigrant communities, who tend to live in over policed areas, have less trust in the police (Wu, Sun, and Triplett 2009) and are thus less likely to initiate calls for service, especially following high profile incidents of police violence (Desmond, Papachristos, and Kirk 2016). As one developer explained to me, using calls for service then reduces over policing in neighborhoods while directing police to neighborhoods who request police presence. Additionally, incident reports for property crimes will be smaller in number in poorer communities who do not carry insurance or do not trust the police to investigate those crimes. How reported crimes intersect with policing practices to produce incident reports and how those reports might be manipulated by police still remains an open question (Ferguson 2017a), but these datasets are easy to defend as a better reflection of crime as it exists than arrest data. If many critiques of predictive policing argue that it reproduces existing policing practices baked into data, then the use of incident reports and calls for service demands different questions to be asked about what biases are present in the data and how they affect policing practices.

If we accept the premise that incident reports and calls for service will lead to policing in communities that request that police presence while reducing policing in communities that do not, then it becomes important to investigate how that plays out in practice. This imagined shift of police resources relies on an understanding of strict and static boundaries between communities that are disproportionately impacted by policing and those who are not. In the diverse and rapidly gentrifying urban areas of contemporary US cities, however, such boundaries are always shifting and already blurry, constantly undermined by patterns of mobility that defy static understandings of who is present in any particular space (Shelton, Poorthuis, and Zook 2015). By redirecting police resources while not being able to account for the racial biases of policing practices, predictive policing might just result in the over policing of the same people but in different spaces. For those who are deemed out of place, especially as they inhabit the blurry borders of race and class separating changing and diverse neighborhoods, predictive policing might just amplify those calls for service that reflect the racist spatial imaginaries of newcomers—the largely affluent, white gentrifiers of urban centers.

Built to facilitate social networking for neighbors, the website Nextdoor reveals how racial profiling operates to increase police presence in particular neighborhoods. People of color regularly become labeled as "suspicious" by white users, who will share info on neighborhood threads about people who they think seem out of place. Based on these suspicions, neighbors will convince each other to call the police, who sometimes have partnerships with the website (Levin 2015; Medina 2018). While Nextdoor did implement algorithmic means to address the problem after it became widely reported, racial profiling, rather unsurprisingly, still persists on the site (O'Donovan 2017). Once called, police are obligated to investigate concerns, file incident reports, and be present in that neighborhood. Data from those call and reports then become part of the data used in predictive policing, leading to the potential for future predicted crimes in those areas. In some cases, neighbors calling the police is an intentional strategy to increase police presence in the neighborhood over time. If the crimes are reported that match up with those used in software, the desire for police presence will be realized through predictive means as software used for policing becomes weaponized by those in privileged positions.

While recognizing the problems with using arrest data as a proxy for crime is certainly important, other types of crime data may only shift policing problems to other geographies. Just as incident reports and calls for service might be proxies for other unaccounted social and economic processes, so too might the correlations produced through models be reflective of something quite different than crime. Following Shannon Mattern's (2018) suggestion in regards to machine learning aimed at public health, correlations demand to be theorized and historicized to understand their mechanisms and causes. Software developers generally recognize that their software does not determine causal mechanisms, regularly offloading that task to crime analysts. But it seems unlikely that the social, historical, and political relations that go into data and its relations can ever be fully explored within the context of policing.

For the makers of such software, theories of data, policing, and processes generally remain limited to criminology literature, which is an exceedingly narrow lens through which to understand policing. In machine learning software, criminology theories will be used to justify engineered features, but developers will regularly be surprised by correlations produced by the machine. Some insights will be explained through reasoning by developers used to justify correlations. For example, if wind speeds correlate with lower assaults, developers will reason that wind makes it unpleasant to be outside, thus reducing the risk of crime. Or if hot days correlate with burglary, they will reason that people leave their windows open when it's hot, leaving them vulnerable to burglars. This process of rationalization is a mode of social theorizing produced by developers in concert with the machines they have programmed, enrolling data, which itself has been produced through vast networks of social and technical relations. When correlations are not explainable through existing criminology literature or modes of reasoning, they are sometimes explained by police who better understand underlying phenomena being exposed by software. For developers, this is a good sign since software is revealing something that is perceived to be true. Always tacking back and forth between criminology theories and intuition of both officers and developers, predictive policing is a process of experimentation with algorithms, data, and thresholds as developers constantly modulate their models, producing new representations of phenomena.

Experimentation is often posed as a necessary component of predictive systems (Shapiro 2018) as they not only attempt to forecast crimes, but also to link forecasts to police practices (Perry et al. 2013). Central to the need for experimentation are the uncertain translations that accompany predictive policing, as described above. Glimpses of how data does or does not reflect the reality it attempts to measures, how forecasts are used as actionable evidence, how officers use and understand software, and how software affects crime rates can all be gleaned, in part, through developer tweaking algorithms, trying new data, implementing new protocols, and seeing how things change. With growing public concern over the use of software for surveillance and policing comes growing desire to understand these behind-the-scenes processes.

Translation IV: Between developers and the public

Developers perceive a void between their work and public understandings of predictive policing fueled by sensationalist reporting. My initial contact with a software company working on predictive policing was met with some wariness as they expressed that they wanted to avoid another "hot take" that misrepresented what they were doing. Critiques of software in the media, in the eyes of developers, often profoundly misunderstand both technology and policing. For example, a widespread critique correctly identifies the need to understand how data is produced and used in models, but incorrectly claims that predictive policing uses arrest data to model crime. This line of reasoning argues that arrest data represents officer bias, which then becomes sedimented within models, leading to feedback loops that reinscribe those biases, leading to more arrests of the same people in the same places. Despite the fact that there are no known vendors using arrest data for place-based predictive policing, this critique is pervasive in the media, leading to developers expressing their frustration and withdrawing from public debate. As one developer told me, the narrative pushed by media, both liberal and conservative, forecloses the possibility for nuanced and well-informed debates that actually address the real problems of both policing and predictive software.

Many critiques draw from popular culture, envisioning policing's descent into technologically-augmented dystopian fantasies. In countless interviews I have conducted and talks I have watched, developers make a point to distinguish their work from that imagined by the movie *Minority Report*. Since endless media reports describe predictive policing as a way to predict and prevent future crimes before they happen, *Minority Report*, in which police are able to arrest people before they even commit a crime, becomes a convenient grounding example. Developers of place-based policing, in the US context where predictive policing has become a topic of public debate, are careful to point out the differences between their work of crime forecasting and the more dystopian vision of being arrested for pre-crimes. Despite their wariness of such comparisons, the future of policing as depicted in science fiction does inform some ambivalence on the part of developers. As one developer said to me, "I mean I'm not fully on board with a police state monitoring everything, drones everywhere… you have to be really careful and… try to implement a system that is not going to put people's civil liberties in jeopardy."

In describing predictive policing, developers will also use critiques as a means of contrast against how things actually work. For example, by citing concerns over government surveillance or biased policing, developers can show how their methods circumvent their concerns. Critiques, then, become useful illustrations of the dangers of doing predictive policing the wrong way, thus producing strong arguments to potential clients as to why the adoption of particular systems are not subject to the same concerns guiding public debate. This is not to say that such claims are empty, but rather, that critique shapes how technologies are developed and understood, even if they are not always accurate. In evidence-based policing forums in particular, responses to critiques guides discussions about new technologies. Of course, just like there is resistance to new policing technologies, there is also resistance that disregards critiques as uniformed of the needs, demands, and difficulties of policing. Additionally, those software vendors who do engage with critiques are a self-selecting group—many vendors hide their methods, algorithms, and strategies from public or academic scrutiny.

Transparency also plays a formative role in how the public and developers view technologies. Near-repeat models, for example, are relatively easy to explain based on theories of repeat victimization. Additionally, they have been written about extensively in academic literature, which theorizes them in relation to multiple studies (e.g., Bowers et al. 2004). In contrast, machine learning approaches are much more complex, relying on the machine to find complex patterns of correlation previously unknown to researchers. As one developer explained, machine learning approaches can be a hard sell to skeptical officers since transparency is lost to the complexity of the algorithms. In some cases, machine learning algorithms can reveal the variables that went into making a particular prediction, allowing for some transparency. Often, however, like in the case of neural network models, the factors that went into making a particular prediction can be opaque even to developers[41]. Lack of transparency is also a concern for how predictive software gets used as there seem to be no public policy guides for how the software should be used in departments (Robinson and Koepke 2016). So while there are some software vendors that participate in public and academic forums, opening their methods and practices to scrutiny, lack of transparency is still a widespread problem that feeds public imaginaries of predictive policing. The task ahead, and part of the goals of this chapter, is to open critiques of predictive policing to specific concerns and problematics that get at the heart of how predictive

[41] 41. Limited is known about many machine learning approaches used in predictive policing software, but there are no known packages currently using neural network approaches.

policing software is understood, produced, and unevenly integrated into existing infrastructures of policing.

Cultivating doubt

Uncertain translations produce doubt amongst predictive policing developers, which sometimes comes to the fore in conversations. Recognizing the limits of data, statistics, and RCT results, developers waiver between touting their work and recognizing the limits of what can actually be claimed. How technologies integrate into and change networks of practice is always difficult to ascertain as the use and understanding of technology often conflicts with its planned intent (see for example: Suchman 1985; Pinch and Bijker 1987). These difficulties are compounded by structural factors unique to policing, as described above, which often lead to police departments resisting outside evaluations. So while developers can easily point to how well a particular model fits a given dataset, the uncertain translation between predictions and practice introduces doubt into the enterprise of software development.

Developers will defend their particular model choices while recognizing the limits of endlessly refining them. Once a certain threshold of predictive accuracy has been reached, developers recognize the diminishing returns of working on the technology alone. Instead, how those predictions translate into practice becomes paramount to successful implementation strategies. “It makes almost no difference which algorithm you use,” one developer explained to me, “what matters is how the police turn [predictions] into actionable plans.” Another developer echoed that sentiment, observing that, “there is an increasing focus on accuracy and I think it's in some ways misplaced. As long as you have a decent level of accuracy, I think incremental changes don't really matter too much.” The recognition of the limits of their algorithms leads developers to think seriously about how they integrate with existing practices of policing, whether that be through appeals to officer psychology, exploitation of the hierarchical structure of police departments, or through experimentation with functions and features.

Even if police officers used the software as imagined by developers, doubts would still persist for some. As one developer told me, "as far as I'm concerned… it's a completely unknown question about whether police patrol actually prevents crime, in any sense other than you're literally standing there when a crime is going to happen and you prevent that." While some commercial software advertisements might imply that predictive policing will lead officers to crimes in progress, this developer told me that this is almost never the case. Instead, predictive policing relies on the belief in deterrence effects, as described earlier, even if the evidence in support of such effects are limited. The translation between predictions and practice is partially blocked by the blue wall of policing and additionally muddied by the situated contexts and contingencies that mark the complex socio-technical network of predictive policing. From this uncertainty arises doubt about how predictive policing software is related to crime, in terms of how crime is modeled and if it can effectively be prevented.

Uncertainty, however, does not negate the role that developers play in coding and recoding people and places as being risky. There is always a politics to classification, as categories reflect particular beliefs and subsequently shape how social systems are understood (Bowker and Star 1999). As described earlier, model choices are reflective of ideological choices on the part of developers. Imagining crime as a contagious disease requiring treatment, for example, is one way of narrating crime theory. Another is based on optimal forager theories borrowed from ecology, as criminals are likened to animals foraging for food. In RTM models, the focus is on the environmental factors that determine crimes, marking some areas as inherently dangerous due to their geographies. And machine learning attempts to incorporate any and all data and theories, allowing the computer to automatically find the most salient correlations. How these understandings of crime filter into discourse, classification schemes, police practices, and popular media all have important effects on how crime is produced as a concept and addressed as a society.

Place-based predictive policing, if successfully implemented, will shift the geographies of policing by coding particular places as risky during particular times of the day or year. The uncertain translations that accompany its implementation have important implications for the people who inhabit those places. Uncertainty becomes the grounds for asking specific questions about how algorithms touch down in particular places, linking to existing structures, practices, and networks of relations. In the end, it might just turn out that Andrew Ferguson's observation that "[t]he problem with predictive policing is the policing part" might just be true (quoted in Lorinc 2018). If so, the insights we gain from computational approaches to predicting crime might just lead us to other causes, correlations, and ways of thinking about the machines we build.

Conclusion

In one of Harun Farocki's (1969) early films—the semi-fictional, fake documentary *Inextinguishable Fire*[42]—the viewer encounters the infrastructure of the war machine. In the mundane site of power of the factory, where chemicals are produced and combined to make napalm for the war, scientists and workers have ambivalent positions in relation to their contribution to killing people in Vietnam. Some justify the need for killing by claiming it will ultimately save lives, others argue that the chemicals they make have many other necessary uses, while others do not even recognize the connection. In one scene, a man in a suit, sitting in a Dow

[42] Thanks to Brian Holmes for pointing out the resonances between this research and Farocki's film.

office defends the war, while justifying Dow's contribution by observing, "The State Department has given us millions for the further development of napalm. Parts of the public, as well as some of our employees don't understand this. A chemical corporation is like a set of building blocks. We let each worker have one block to work on. The we put the blocks together to make whatever our clients request."

Similarly, in interchangeable software officers, mostly young people (and in the case of predictive policing: mostly young men) build pieces of software that increasingly inform political decision-making in the 21st century. Built with lofty ideals of social good, the methods and machines combine in surprising ways as they integrate into existing socio-technical networks, produce and modulate the spaces of everyday life, and become part of sedimented practices and ways of knowing the world. It is too early to know exactly what is being built, but tendencies are beginning to emerge that give us glimpses of what that future might look like, which in many cases is just the acceleration and sedimentation of the past. How the pieces go together to build something bigger is always being shaped and reshaped.

After the chemical plant, Farocki takes us to a vacuum cleaner factory, where one worker slowly steals pieces to take home, hoping that he can assemble his own vacuum cleaner at home. But each time he does so, he ends up with a submachine gun. A student working in the same vacuum factory is convinced that they are actually building submachine guns, so he slowly steals pieces to take home. But each time he assembles them, he ends up with a vacuum cleaner. The film ends with an engineer observing, "I'm an engineer and I work for an electrical corporation. The workers think we're making vacuum cleaners. The students think we're making submachine guns. This vacuum cleaner can become a useful weapon. This submachine gun can become a useful household gadget. What we manufacture depends on the workers, students and engineers."

5. Conclusion: future directions

In the three case studies that make up this dissertation, I deploy a mixed method, qualitative approach to studying the digital geographies of governance. In Chapter 2, I examine how infrastructures change when algorithms are integrated into them. Drawing from recent reports of algorithmic deception and contestation, theories of infrastructure, and software studies, I argue that infrastructures tend to become reconfigurable, interoperable, and deceptive once controlled by computers. These tendencies, I conclude, have important implications for how spaces and subjects are constructed by automated and often opaque processes. Using an archival approach, in Chapter 3, I show how software affords new possibilities for surveillance. Through an examination of the crowdsourced effort to identify the culprits of the 2013 Boston Marathon bombing, I show how disparate systems were brought together to surveil individuals, construct complex understandings of space, and influence the event as it transpired. And finally, in Chapter 4, I use interviews, workplace and observation, and document analyses to study the algorithms of predictive policing. In that chapter, I show how software developers negotiate the challenges that come with modeling crime. I argue that uncertain translations undergird their work, producing doubt amongst developers, but also opening critical questions that can help reveal how software changes practice of policing. In what follows, I make three methodological suggestions for the field of digital geographies. All three have informed my recent work and point to future research directions as I continue to grapple with the integration of algorithms into processes of governance.

Three methods for digital geographies

With the recent interest in digital geographies[43], there have been a number of methodological and conceptual suggestions for researching the digital (see, for example: Rose 2016; Dalton, Taylor, and Thatcher 2016; Kitchin, Lauriault, and Wilson 2017; Leszczynski 2017). In what follows, I suggest three methodological directions for digital geographies informed by the research outlined in the previous chapters. They include experimental probing of computational systems, reading literature in computer science and adjacent disciplines, and building software as a means to show other ways of knowing the world through computation. These suggestions are meant to supplement, not replace, the existing diversity of methods in digital geographies, while recognizing that some geographers are currently deploying these methods.

First, following Matthew Fuller's (2007) suggestion that "the only way to understand how complex media objects interact is to carry out those interactions" (1), I think there is an important place for experimental digital methods to probe software and algorithmic infrastructures to explore their shape, relations, and functions.[44] Since so much of computing exceeds our ability to perceive or grasp it, digital tools have the potential to explore the materiality of algorithmic infrastructures by intervening in them at the level of their operations. Consider, for example, the Shodan browser[45], which works by probing random IP addresses to find internet-connected devices. It has exposed and mapped a vast networks of computers connected to the public internet, including everything from software controlling critical infrastructures like power plants to internet-connected webcams in people’s homes. Not even Google, who actively contributes to

[43] In the last year, Digital Geographies groups have formed as part of the Royal Geographical Society with the Institute of British Geographers (RGS-IBG) and the American Association of Geographers (AAG), two of the major English-language academic geography associations.

[44] This suggestion has resonance with geographic approaches suggested by work in site ontology. For example, Woodward et al. (2010) make an analogy between research and the game of pick-up sticks, suggesting a process of experimentation where one tests the relations, pressures, and intensities of the sticks in their various relations (276).

[45] See: https://www.shodan.io/

the structure of the internet by making connections possible, can know everything about its structure. They too must use novel methods like web crawlers and PageRank algorithms to reverse-engineer and "read" the infrastructure of the internet (Peters 2015). Building our own digital tools can similarly work towards studying the complexities of algorithmic systems, revealing a partial view that can contribute to efforts to produce rich, complex, or speculative accounts of geographies entangled with digital processes.

Geography is particularly well-positioned to undertake this work owing to its long history of grappling with computational problems (Lally and Burns 2017), its production of critical scholars who use and build digital tools (Schuurman and Pratt 2002; Thatcher et al. 2016), and the interdisciplinary nature of the field, which often includes collaborative, mixed-methods research. Additionally, the public release of exploratory digital tools—for example, one can download directions for building cell-site simulator detectors like those used secretly by police departments (Ney et al. 2017)— can act as starting points for this type of research (See also: DaCosta 2017). Informed by spatial questions related to computation, custom digital tools have the potential to open geographic research to new ways of understanding how the complex and shifting material relations of algorithmic infrastructures contribute to how space is understood, produced, and remade. How these material configurations construct the world is both complex and always in flux as algorithmic infrastructures constantly shift (reconfigurability), always bringing new phenomena and data into the fold (interoperability), while hiding their operational logics and functions (deception).

Second, we can learn a lot about current tendencies and possibilities for computation by reading computer science, data science, and adjacent fields that both explore and build emerging systems. The findings reported in these literatures give important insights into current methods for sorting, analyzing, and representing data (e.g., de Montjoye et al. 2013; Kosinski, Stillwell, and Graepel 2013); emerging computational possibilities (e.g., Wu et al. 2012; Ali et al. 2015); and methods for exploring, reverse-engineering, and intervening in existing software (e.g., Diakopoulos 2014). These insights can deepen geographic research that looks to the specific functioning of computational systems. Additionally, careful analysis of these literatures can help resist the tendency to adopt popular computational imaginaries rooted in buzzwords like "intelligence" and "automation" that often act to obscure material relationships and processes. This literature might also be the starting point for developing projects and collaborations rooted in geographic questions or to provide a socially and spatially-grounded critiques that can complement the conclusions that are drawn from computational studies.

And, finally, geography can participate in the development of other ways of knowing the world through computation, some of which might be informed by the previous two suggestions. An exploratory and speculative approach to computing (Drucker 2009) can reveal latent possibilities for transforming computational systems that subtend everyday life while hinting at alternative paths the development of technologies might have taken. This suggestion is informed by the desire for critique to be a constructive exercise of building, which entails a "care for the subject" of our critique (Schuurman and Pratt 2002, 291) and the recognition of the transformative possibilities for reconfiguring modes of knowledge production—a central lesson of STS. For example, recognizing the limitations of spatial data and GIS software, some geographers have recently proposed data formats (Bergmann 2016) and mapping techniques (O'Sullivan, Bergmann, and Thatcher 2017) that eschew taken-for-granted Euclidean understandings of space in favor of relational understandings. If standardized algorithmic methods for sorting and representing spatial data make possible the digital maps that have become ubiquitous in everyday life, attempts to rethink these methods help us rethink the assumptions of existing systems while bringing other ways of knowing the world to fruition.

These three suggestions are meant to act as starting points for efforts to probe and experiment with algorithms that tend to evade traditional modes of geographic scrutiny. While presented here as discrete and potentially ambitious projects, they might also be used in smaller, discrete moments as a means to inform and guide other modes of inquiry. For example, an ethnographer of software developers might use these suggestions to draw out developers' validation attempts, which are often a digital means of probing the infrastructures they build. Or, inspired by ProPublica's revealing attempts to buy discriminatory advertisements on Facebook (Tobin 2017), the digital geographer might think of other, small ways to poke the systems they study at the level of their operations and see what insights they might kick out. In the next section, I outline a proposal for predictive policing that encompasses these three methodological suggestions.

Reversing polarity: a predictive policing proposal

How predictive policing will be taken up, how it will change policing practices, and the accuracy of the assumptions central to its development all remain open questions, but, in spite of all of this, it might be deployed to better understand existing policing practices. Andrew Ferguson (2017b), for example, argues that the technologies that make predictive policing possible also could be used to collect and analyze what he calls “blue data.” This “blue data,” argues Ferguson, might include officers' locations, arrest data, or even bodycam audio and video, all of which can be analyzed to see how officer patrol routes line up with known crimes distribution, bias rates in arrests, the quality of interactions with the public, or predictions for individual officer violence[46]. Additionally, Jerry Ratcliffe (2015) has very tentatively suggested an approach

[46] See also: https://dsapp.uchicago.edu/projects/public-safety/early-warning-and-intervention-systems-for-police-departments/

to measuring harm that would include certain types of officer interactions, including stops and arrests, as themselves constituting harm to a community. One developer I spoke with also outlined an initial proposal to include harms of policing within predictive policing models, mostly focusing on reducing repeated stops of people. Including harm in model calculations is not unusual, but it has generally focused on the idea of harm reduction through crime reduction. In other words, preventing different types of crimes leads to different levels of harm reduction within a community. Recognizing policing as a harm itself would represent a marked shift in how policing is typically understood within departments. The proposals above are all rather cautious in their approach, in part, I am sure, any kind of move in that direction will likely be politically unworkable in most departments. In spite of this, they begin to recognize the disparate impact of policing and the community resentment that it can precipitate. One need only remember the damning report outlining the abuses of the Ferguson Police Department against the community in the wake of Michael Brown's murder by the police ("Investigation of the Ferguson Police Department" 2015). But instead of cautious approach to studying the harms of policing, what if we followed this suggestion much further through to its logical conclusion and produce a more radical vision of what an understanding of the harms of policing might entail?

In what follows, I sketch out some possible directions for measuring harm using predictive policing methods and technologies as described in Chapter 4. The proposed methods are informed by problems recognized by researchers and developers regarding policing in practice and provide methods for predictive policing to address structural determinants of crime more substantively than current systems. The suggestions that follow may also serve as starting points for research into the disparate impacts of policing in general or the specific impacts of predictive policing. Taken as a trajectory for future research, this proposal adopts the three methodological suggestions outlined above. Specifically, it requires experimentation with current systems of predictive policing informed by computer and data science literature with the goal of developing prototypes that reimagine the software in question. Instead of an additive approach, these suggestions use data to subtract practices, places, and methods from policing, thus reversing the polarity of how data-driven policing is usually implemented.

First, if predictive policing is conceptualized as a means of crime deterrence, the ability for officers to arrest and use force against people undermines the ability to test whether or not this is true. Additionally, evidence suggests that sending officers to places where increased crime risks are forecasted results in increased arrest rates in those places (Brantingham, Valasik, and Mohler 2018). In some implementations of predictive policing in the UK, unarmed, non-arresting officers, bearing the title of "police community support officers" (PSCO) will be dispatched on foot patrols to deter predicted crime. As one developer told me, they will also send other public workers who wear high-vis jackets through risky areas as a means to deter crime. There is some evidence that unarmed and non-arresting officers can reduce crime rates in the UK owing to their visibility alone (Ariel, Weinborn, and Sherman 2016). Similar studies in the US could verify claims of deterrence, minimize over policing of particular people and neighborhoods, and work towards a de-escalation of police violence. In situations where predictive software forecasts property crimes like burglary, official, non-deputized presence as a deterrence mechanism can easily be tested with minimal risk. Using data from predictive policing systems, researchers could compare current distributions of arrest and risk to forecast how the deployment of non-arresting officers would change outcomes, thus producing a strong argument for trials that include non-arresting officers. Additionally, this suggestion can serve as a starting point for research into how different policing tactics in conjunction with crime forecasts correlate with arrest, search, and citation rates.

Second, existing distributions of arrest can be used in combination with other data to measure existing policing bias, which can then be used as a metric in current predictions. For example, in the Lum and Isaac (2016) study mentioned earlier, drug crime predictions using predictive policing trained on arrest data were compared to best estimates of actual distributions of drug crimes. Using the difference between the two to ascertain where biases are most pronounced, developers could compare this arrest bias with current predictions. Biased areas can then be subtracted from the model to to avoid adding police presence to already over policed areas where people are disproportionately arrested. This suggestion additionally recognizes that predictive policing only directs officers to forecasted areas when they are not otherwise engaged in responding to calls or doing prescribed actions. The subtraction of already over policed areas will merely redistribute them at times when they are not already engaged with other activities, which will likely often be in those areas experiencing heightened arrest rates. This suggestion could also be used as a means to study existing practices and predictions to developer deeper critiques of predictive systems.

Third, building on the second suggestion, predictive policing can take long term community harm into account when distributing officers and subtract those areas from officer dispatches. Michelle Alexander (2012) argues that mass incarceration has produced a racialized caste system in the US that denies a large portion of African Americans access to social and economic opportunities. As Alexander shows, over half of young adult black men in some major cities are under the control of the justice system, which includes those in prison or jail and under parole or probation. The long term, community impact of these processes have been devastating in places, while the consequences of otherwise minor arrests can be huge for those on probation. Accounting for the unequal and long term harm that communities and individual suffer under routine policing can be another factor in studying existing systems as well as in determining police distributions and rethinking police tactics, as outlined in the next suggestion.

Fourth, and finally, both RTM and machine learning approaches to predictive policing locate specific factors that contribute to crime. The producers of RTM go so far as to argue that they add the *why* to the *when* and *where* of predictive systems. If this is true, and here I follow the public health metaphors that are so ubiquitous in criminology, then it makes no sense—economically, socially, or scientifically—to continue only treating the symptoms of crime with dosages of police. Instead, predictive systems can be used to ascertain correlations—both positive and negative —between particular geographic features or social processes and crime and work towards remedying problems or making better conditions for the reduction of crime (Ferguson 2017b). Following Ferguson (2017b), "If environmental vulnerabilities encourage crime, changing the environment might matter more than policing the area" (80). The logical outcome of such a program is probably the least politically viable at the moment as it entails a radical and structural critique of existing social and economic conditions, requiring more far reaching and fundamental changes to implement than can be dreamt of within the bounds of police reform. But using this approach can be the starting point for experimenting with other data and computational systems as a means to undermine the conclusion that policing is the right answer to crime.

Conclusion

The above proposal takes advantage of the reconfigurability of algorithmic infrastructures as a means to research them, while showing other ways that governance could be coded. By appropriating and reconfiguring the data, methods, and techniques that make predictive policing possible, this proposal seeks to show other ways that policing and crime can be conceptualized, understood, and subject to critique. Building from the uncertainties of translations observed in fieldwork, the proposal illustrates how qualitative research can contribute to interventions that engage with the technical apparatuses of study. Integrating critical practices of research with critical modes of making is one possible direction forward for digital geography scholarship. While this dissertation focuses on qualitative study, this conclusion suggests ways for case studies to develop into critical making practices, highlighting how software interventions can contribute other modes of understanding the ramifications of algorithmic forms of governance.

Software interventions integrate with the three meanings of coded outlined in the introduction, which refer to coded as (1) computer code used to build software, (2) the black-boxed tendencies of computational systems, and (3) the automated classification of people and places. First, since some of the computational methods that make up software systems are known or can be inferred, analogous systems can be created that reconfigure existing systems based on other conceptual and theoretical frameworks. In other cases, where little is known about the software in question, identical or similar datasets can be analyzed in ways that highlight computational transparency and how decisions impact the insights that software makes possible. Second, building software can raise important questions about existing black-boxed systems. For example, if the insights provided by an opaque system are found to be consistently racially biased (e.g., Angwin et al. 2016), software interventions can provide alternative ways of conceptualizing social problems. The proposal above, for example, recombines existing policing software with other social indicators to account for long term, structural processes that contribute to crime in ways that current systems do not take into account. These alternative prototypes can be the starting point for asking important questions about the assumptions, methods, and data that go into black-boxed systems. And third, software interventions can rethink how people and places become classified and coded, or even undermine the need for discrete and static categorization schemes. The simplification of complex, fluid, and intertwined social processes to the limited categorical schema of maps, which I showed in late 19th century in the introduction, is a narrow lens through which to see the world and inform political decisions. While software used for political decision-making often results in limited categorization schemes, software also affords possibilities for developing relational, processual, and dynamic modes of visualization that would reflect other political orientations and spatial imaginaries. Taken together, the three understandings of "coded governance" highlighted above contribute to the themes running through this dissertation while pointing to other ways of conceptualizing software as an object of study.

The difficulties of studying software, many of which have been outlined in the preceding chapters, point to the need for a variety of methodological and theoretical orientations that can produce rich and complex accounts of coded governance. Only grasped in partial glimpses, software often evades scrutiny, but expanding our critical tools has become important to understanding the implications of a world increasingly governed in concert with machines. The suggestions outlined above point towards ways to both critically interrogate but also intervene in software systems, opening up new possibilities for both studying software and seeing the world through computation.

References

Aaronson, Daniel, Daniel Hartley, and Bhashkar Mazumder. 2017. “The Effects of the 1930s HOLC ‘Redlining’ Map;” *Federal Reserve Bank of Chicago Working Paper* 2017–12.

Adegoke, Yemisi. 2017. “Uber Drivers in Lagos, Nigeria Use Fake Lockito App to Boost Fares — Quartz.” *Quartz Africa*. November 13, 2017. https://qz.com/1127853/uber-drivers-in-lagos-nigeria-use-fake-lockito-app-to-boost-fares.

Adey, Peter. 2009. “Facing Airport Security: Affect, Biopolitics, and the Preemptive Securitisation of the Mobile Body.” *Environment and Planning D: Society and Space* 27 (2): 274–95. doi:10.1068/d0208.

Ahmed, Maha. 2018. “Aided by Palantir, the LAPD Uses Predictive Policing to Monitor Specific People and Neighborhoods.” *The Intercept*. May 11. https://theintercept.com/2018/05/11/predictive-policing-surveillance-los-angeles/.

Alexa. 2016. "reddit.com Site Overview." Accessed on February 9, 2016. http://www.alexa.com/siteinfo/reddit.com

Alexander, Michelle. 2012. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. New York: New Press.

Ali, Kamran, Alex Xiao Liu, Wei Wang, and Muhammad Shahzad. 2015. “Keystroke Recognition Using WiFi Signals.” In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking - MobiCom ’15*, 90–102. Paris, France: ACM Press. doi:10.1145/2789168.2790109.

Alpaydin, Ethem. 2016. *Machine Learning: The New AI*. Cambridge, MA: MIT Press.

Althusser, Louis. 2001. “Ideology and Ideological State Apparatuses.” In *Lenin and Philosophy, and Other Essays*. New York: Monthly Review Press.

Amoore, Louise. 2011. “Data Derivatives: On the Emergence of a Security Risk Calculus for Our Times.” *Theory, Culture & Society* 28 (6): 24–43. doi:10.1177/0263276411417430.

———. 2013. *The Politics of Possibility: Risk and Security beyond Probability*. Durham: Duke University Press.

———. 2016. “Cloud Geographies: Computing, Data, Sovereignty.” *Progress in Human Geography*. doi:10.1177/0309132516662147.

———. 2017. “What Does It Mean To Govern With Algorithms?” *AntipodeFoundation.org*. May 19, 2017. https://radicalantipode.files.wordpress.com/2017/05/2-louise-amoore.pdf.

Amoore, Louise, and Marieke De Goede. 2005. “Governance, Risk and Dataveillance in the War on Terror.” *Crime, Law and Social Change* 43 (2–3): 149–73. doi:10.1007/s10611-005-1717-8.

Amoore, Louise, and Rita Raley. 2017. “Securing with Algorithms: Knowledge, Decision, Sovereignty.” *Security Dialogue* 48 (1): 3–10. doi:10.1177/0967010616680753.

Anderson, Ben. 2010. “Preemption, Precaution, Preparedness: Anticipatory Action and Future Geographies.” *Progress in Human Geography* 34 (6): 777–98. doi:10.1177/0309132510362600.

Anderson, Chris. 2008. "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete." *WIRED*. June 23, 2008. https://www.wired.com/2008/06/pb-theory/.

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

Angwin, Julia, Jeff Larson, Surya Mattu, Lauren Kirchner, and ProPublica. 2016. "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks. " *ProPublica*. May 23. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Ariel, Barak, Cristobal Weinborn, and Lawrence W Sherman. 2016. "'Soft' Policing at Hot Spots—Do Police Community Support Officers Work? A Randomized Controlled Trial." *Journal of Experimental Criminology* 12 (3): 277–317. doi:10.1007/s11292-016-9260-4.

Arkell, George E., and Ernest Aves. 1899. *Police District 2, District 23, District 24, District 25, District 27, District 28*. https://booth.lse.ac.uk/notebooks/b360

Arthur, Charles. 2011. "Rogue Web Certificate Could Have Been Used to Attack Iran Dissidents." *The Guardian*, August 30, 2011. http://www.theguardian.com/technology/2011/aug/30/faked-web-certificate-iran-dissidents.

Ash, James, Rob Kitchin, and Agnieszka Leszczynski. 2018. "Digital Turn, Digital Geographies?" *Progress in Human Geography* 42 (1): 25–43. doi:10.1177/0309132516664800.

Bachelard, Gaston. 2002. *The Formation of the Scientific Mind: A Contribution to a Psychoanalysis of Objective Knowledge*. Manchester: Clinamen.

Barad, Karen. 2007. *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Durham: Duke University Press.

Barocas, Solon, and Andrew D. Selbst. "Big data's disparate impact." *California Law Review* 104 (2016): 671. doi: 10.15779/Z38BG31

Barnes, Trevor J. 2004. "Placing Ideas: Genius Loci, Heterotopia and Geography's Quantitative Revolution." *Progress in Human Geography* 28 (5): 565–595.

Beer, David. 2009. "Power through the Algorithm? Participatory Web Cultures and the Technological Unconscious." *New Media & Society* 11 (6): 985–1002. doi:10.1177/1461444809336551.

Benjamin, Walter. 1969. "Paris: Capital of the Nineteenth Century." *Perspecta* 12: 163. doi:10.2307/1566965.

Bergmann, Luke. 2016. "Toward Speculative Data: 'Geographic Information' for Situated Knowledges, Vibrant Matter, and Relational Spaces." *Environment and Planning D: Society and Space* 34 (6): 971–89. doi:10.1177/0263775816665118.

Berlant, Lauren. 2016. "The Commons: Infrastructures for Troubling Times*." *Environment and Planning D: Society and Space* 34 (3): 393–419. doi:10.1177/0263775816645989.

Bhardwaj, Anshu. 2014. "Harnessing the Crowd for Neurology Research." *Science Translational Medicine* 6 (250): 250ec141-250ec141. doi:10.1126/scitranslmed.3010124.

Bijker, Wiebe E. 1995. *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change*. Cambridge, MA: MIT Press.

Booth, Charles. 1889. *Labour and Life of the People*. https://archive.org/details/labourlifeofpeop01bootuoft.

Borch, Christian. 2013. *The Politics of Crowds: An Alternative History of Sociology*. Cambridge: Cambridge University Press.

Bowers, Kate J., Shane D. Johnson, and Ken Pease. 2004. "Prospective Hot-Spotting: The Future of Crime Mapping?" *British Journal of Criminology* 44 (5): 641–58. doi:10.1093/bjc/azh036.

Bowker, Geoffrey C., Karen Baker, Florence Millerand, and David Ribes. 2009. "Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment." In *International Handbook of Internet Research*, edited by Jeremy Hunsinger, Lisbeth Klastrup, and Matthew Allen, 97–117. Dordrecht: Springer Netherlands.

Bowker, Geoffrey C., and Susan Leigh Star. 1999. *Sorting Things out: Classification and Its Consequences*. Inside Technology. Cambridge, MA: MIT Press.

boyd, danah. 2011. "Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications." In *A Networked Self: Identity, Community and Culture on Social Network Sites*, edited by Zizi Papacharissi, 39–58. New York: Routledge.

Boyne, Roy. 2001. *Subject, Society, and Culture*. Theory, Culture & Society. London: SAGE.

Brabham, Daren C. 2013. *Crowdsourcing*. Cambridge, MA: The MIT Press.

Brantingham, P. Jeffrey, Matthew Valasik, and George O. Mohler. 2018. "Does Predictive Policing Lead to Biased Arrests? Results From a Randomized Controlled Trial." *Statistics and Public Policy* 5 (1): 1–6. doi:10.1080/2330443X.2018.1438940.

Bratton, Benjamin H. 2015. *The Stack: On Software and Sovereignty*. Software Studies. Cambridge, MA: MIT Press.

Browne, Simone. 2015. *Dark Matters: On the Surveillance of Blackness*. Durham: Duke University Press.

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

Bucher, Taina. 2012. "Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook." *New Media & Society* 14 (7): 1164–80. doi:10.1177/1461444812440159.

Bujlow, Tomasz, Valentín Carela-Español, Josep Solé-Pareta, and Pere Barlet-Ros. 2015. "Web Tracking: Mechanisms, Implications, and Defenses." *ArXiv Preprint ArXiv:1507.07872*.

Bunge, William. 1960. "Theoretical Geography." PhD diss., University of Washington.

Burrell, Jenna. 2016. "How the Machine 'thinks: Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 3 (1). doi:10.1177/2053951715622512.

Buttimer, Anne. 1976. "Grasping the Dynamism of Lifeworld." *Annals of the Association of American Geographers* 66 (2): 277–292.

Callon, Michel. 1986. "Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St. Brieuc Bay." *The Sociological Review* 32: 196–223.

Carroll, Rory. 2012. "UC Davis pepper-spray officer fired despite being cleared by internal panel." *The Guardian.* August 2, 2012. http://www.theguardian.com/world/2012/aug/02/uc-davis-pepper-spray-officer.

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

Cheney-Lippold, John. 2011. "A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control." *Theory, Culture & Society* 28 (6): 164–81. doi:10.1177/0263276411424420.

Christl, Wolfie. 2017. "Corporate Surveillance in Everyday Life: How Companies Collect, Combine, Analyze, Trade, and Use Personal Data on Billions." Vienna: Cracked Labs. http://crackedlabs.org/dl/CrackedLabs_Christl_CorporateSurveillance.pdf.

Chun, Wendy Hui Kyong. 2008. "On 'Sourcery,' or Code as Fetish." *Configurations* 16 (3): 299–324. doi:10.1353/con.0.0064.

———. 2011. *Programmed Visions: Software and Memory*. Software Studies. Cambridge, MA: MIT Press.

Coleman, E. Gabriella. 2014. *Hacker, Hoaxer, Whistleblower, Spy: The Many Faces of Anonymous*. New York: Verso.

Coleman, Mat. 2016. "State Power in Blue." *Political Geography* 51 (March): 76–86. doi:10.1016/j.polgeo.2016.01.008.

Cope, Meghan, and Sarah Elwood, eds. 2009. *Qualitative GIS: A Mixed Methods Approach*. Thousand Oaks, CA: Sage.

Cox, Joseph. 2017. "Scammers Say They Got Uber to Pay Them With Fake Rides and Drivers." *Motherboard*. January 24, 2017. https://motherboard.vice.com/en_us/article/535zdn/scammers-say-they-got-uber-to-pay-them-with-fake-rides-and-drivers.

Crampton, Jeremy, and Andrea Miller. 2017. "Intervention Symposium: 'Algorithmic Governance.'" *AntipodeFoundation.org*. May 19, 2017. https://antipodefoundation.org/2017/05/19/algorithmic-governance/.

Crawford, Kate. 2016. "Can an Algorithm Be Agonistic? Ten Scenes from Life in Calculated Publics." *Science, Technology & Human Values* 41 (1): 77–92. doi:10.1177/0162243915589635.

DaCosta, Marc. 2017. "How to Explore the Hidden World of Radio Waves All Around You." *Motherboard*. December 18, 2017. https://motherboard.vice.com/en_us/article/59wpmn/how-to-explore-the-hidden-world-of-radio-waves-all-around-you.

Dalton, Craig M, Linnet Taylor, and Jim Thatcher. 2016. "Critical Data Studies: A Dialog on Data and Space." *Big Data & Society* 3 (1): 205395171664834. doi:10.1177/2053951716648346.

Dalton, Craig, and Jim Thatcher. 2014. "Dalton and Thatcher Commentary - 'What Does a Critical Data Studies Look like, and Why Do We Care?'" *Society and Space - Environment and Planning D*. May 12, 2014. http://societyandspace.com/2014/05/19/dalton-and-thatcher-commentary-what-does-a-critical-data-studies-look-like-and-why-do-we-care.

Davies, Toby P, and Steven R Bishop. 2013. "Modelling Patterns of Burglary on Street Networks." *Crime Science* 2 (1): 10. doi:10.1186/2193-7680-2-10.

Davis, Fred D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." *MIS Quarterly* 13 (3): 319. doi:10.2307/249008.

Degeling, Martin, and Bettina Berendt. 2017. "What Is Wrong about Robocops as Consultants? A Technology-Centric Critique of Predictive Policing." *AI & Society*, May. doi:10.1007/s00146-017-0730-7.

Deleuze, Gilles. 1992. "Postscript on the Societies of Control." *October* 59: 3–7.

Desmond, Matthew, Andrew V. Papachristos, and David S. Kirk. 2016. "Police Violence and Citizen Crime Reporting in the Black Community." *American Sociological Review* 81 (5): 857–76. doi:10.1177/0003122416663494.

Descartes, René. (1637) 1984. *Discourse on Method and the Meditations*. Harmondsworth: Penguin Books.

Dewey, Caitlin. 2014. "How an anonymous Twitter sleuth may have solved a Philadelphia hate crime (and restored our faith in the Internet). " *The Washington Post,* September 17, 2014. 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.

Diakopoulos, Nicholas. 2014. "Algorithmic Accountability Reporting: On the Investigation of Black Boxes." *Tow Center for Digital Journalism*. http://towcenter.org/wp-content/uploads/2014/02/78524_Tow-Center-Report-WEB-1.pdf.

Dourish, Paul, and Genevieve Bell. 2007. "The Infrastructure of Experience and the Experience of Infrastructure: Meaning and Structure in Everyday Encounters with Space." *Environment and Planning B: Planning and Design* 34 (3): 414–30. doi:10.1068/b32035t.

Drucker, Johanna. 2009. *SpecLab: Digital Aesthetics and Projects in Speculative Computing*. Chicago: University of Chicago Press.

Dubrofsky, Rachel E., and Shoshana Magnet, eds. 2015. *Feminist Surveillance Studies*. Durham: Duke University Press.

Edwards, Paul N. 1998. "Y2K: Millennial Reflections on Computers as Infrastructure." *History and Technology* 15 (1–2): 7–29. doi:10.1080/07341519808581939.

———. 2010. *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. Cambridge, MA: MIT Press.

Edwards, Paul N., Geoffrey C. Bowker, Steven J. Jackson, and Robin Williams. 2009. "Introduction: An Agenda for Infrastructure Studies." *Journal of the Association for Information Systems* 10 (5): 6.

Eghbal, Nadia. 2016. “Roads and Bridges: The Unseen Labor Behind Our Digital Infrastructure.” *Ford Foundation*. https://www.fordfoundation.org/library/reports-and-studies/roads-and-bridges-the-unseen-labor-behind-our-digital-infrastructure.

Eubanks, Virginia. 2017. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York, NY: St. Martin’s Press.

Farocki, Harun. 1969. *Inextinguishable Fire*. 16mm film, 25 min.

FBI. 2013a. “Statement of Attorney General Eric Holder on the Ongoing Investigation into Explosions in Boston. ” 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.” 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.

Ferguson, Andrew G. 2017a. “Policing Predictive Policing.” *Washington University Law Review* 94 (5): 1109–89.

———. 2017b. *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. New York: New York University Press.

Fleck, Ludwik. 2008. *Genesis and Development of a Scientific Fact*. Sociology of Science. Chicago: University of Chicago Press.

Foord, Jo, and Nicky Gregson. 1986. “Patriarchy: Towards a Reconceptualisation*.” *Antipode* 18 (2): 186–211.

Foucault, Michel. 2009. *Security, territory, population: lectures at the Collège de France, 1977-1978*. New York: Picador/Palgrave Macmillan.

Fuller, Matthew. 2007. *Media Ecologies: Materialist Energies in Art and Technoculture*. Leonardo. Cambridge, MA: MIT Press.

“FTC Announces Internet of Things Challenge to Combat Security Vulnerabilities in Home Devices.” 2017. *Federal Trade Commission*. January 3, 2017. https://www.ftc.gov/news-events/press-releases/2017/01/ftc-announces-internet-things-challenge-combat-security.

Galloway, Alexander R. 2004. *Protocol: How Control Exists after Decentralization*. Cambridge, MA: MIT Press.

Garnier, Simon, Joel M. Caplan, and Leslie W. Kennedy. 2018. “Predicting Dynamical Crime Distribution From Environmental and Social Influences.” *Frontiers in Applied Mathematics and Statistics* 4. doi:10.3389/fams.2018.00013.

Gatrell, Anthony C. 1983. *Distance and Space: A Geographical Perspective*. New York: Clarendon Press.

Gillespie, Tarleton. 2010. “The Politics of ‘Platforms.’” *New Media & Society* 12 (3): 347–64. doi:10.1177/1461444809342738.

Gillespie, Tarleton. 2014. “The Relevance of Algorithms.” In *Media Technologies: Essays on Communication, Materiality, and Society*, edited by Pablo J Boczkowski, Kirsten A Foot, and Tarleton Gillespie, 167–94. Cambridge, MA: MIT Press.

Glinton, Sonari. 2015. “How A Little Lab In West Virginia Caught Volkswagen’s Big Cheat.” *NPR*. September 24, 2015. https://www.npr.org/2015/09/24/443053672/how-a-little-lab-in-west-virginia-caught-volkswagens-big-cheat.

Goffey Andrew. 2008. “Algorithm.” In *Software studies: a lexicon*, edited by Matthew Fuller, 15-20 Cambridge, MA: MIT Press.

Goodin, Dan. 2017. “Cryptojacking Craze That Drains Your CPU Now Done by 2,500 Sites.” *Ars Technica*. November 8, 2017. https://arstechnica.com/information-technology/2017/11/drive-by-cryptomining-that-drains-cpus-picks-up-steam-with-aid-of-2500-sites/.

Graham, Stephen DN. 2005. “Software-Sorted Geographies.” *Progress in Human Geography* 29 (5): 562–580.

Gregory, D. 2011. “From a View to a Kill: Drones and Late Modern War.” *Theory, Culture & Society* 28 (7–8): 188–215. doi:10.1177/0263276411423027.

Grusin, Richard A. 2010. *Premediation: Affect and Mediality after 9/11*. New York: Palgrave Macmillan.

Haggerty, Kevin D. 2006 “Tear down the walls: On demolishing the panopticon.” In *Theorizing Surveillance: The Panopticon and Beyond*, edited by David Lyon, 23-45. New York: Routledge.

Haraway, Donna. 1988. “Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective.” *Feminist Studies* 14 (3): 575. doi:10.2307/3178066.

———. 1991. *Simians, Cyborgs, and Women: The Reinvention of Nature*. New York: Routledge.

Harley, J B. 1989. “Deconstructing the Map.” *Cartographica: The International Journal for Geographic Information and Geovisualization* 26 (2): 1–20. doi:10.3138/E635-7827-1757-9T53.

Harvey, David. (1973) 1988. *Social Justice and the City*. Oxford: Basil Blackwell.

Hayles, Katherine. 1999. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: University of Chicago Press.

———. 2005. *My Mother Was a Computer: Digital Subjects and Literary Texts*. Chicago: University of Chicago Press.

Hillier, Amy E. 2003. “Redlining and the Home Owners’ Loan Corporation.” *Journal of Urban History* 29 (4): 394–420. doi:10.1177/0096144203029004002.

Hotten, Russell. 2017. “Volkswagen: The Scandal Explained” *BBC News*. December 10, 2017. http://www.bbc.com/news/business-34324772.

Howe, Cymene., Jessica Lockrem, Hannah Appel, Edward Hackett, Dominic Boyer, Randal Hall, Matthew Schneider-Mayerson, et al. 2016. “Paradoxical Infrastructures: Ruins, Retrofit, and Risk.” *Science, Technology & Human Values* 41 (3): 547–65. doi:10.1177/0162243915620017.

Howe, Jeff. 2006. “The Rise of Crowdsourcing.” *Wired*, June 1, 2006. https://www.wired.com/2006/06/crowds/.

———. *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*. 1st paperback ed. New York: Three Rivers Press.

Hunt, Priscilla, Jessica M. Saunders, and John S. Hollywood. 2014. *Evaluation of the Shreveport Predictive Policing Experiment*. Santa Monica, CA: RAND Corporation.

International Association of Chiefs of Police. 2015. “2015 Social Media Survey Results.” http://www.iacpsocialmedia.org/wp-content/uploads/2017/01/FULL-2015-Social-Media-Survey-Results.compressed.pdf.

Introna, Lucas D. 2011. “The Enframing of Code: Agency, Originality and the Plagiarist.” *Theory, Culture & Society* 28 (6): 113–41. doi:10.1177/0263276411418131.

Isaac, Mike. 2017. “How Uber Deceives the Authorities Worldwide.” *The New York Times*, March 3, 2017. https://www.nytimes.com/2017/03/03/technology/uber-greyball-program-evade-authorities.html.

Isaac, William, and Kristian Lum. 2018. “Setting the Record Straight on Predictive Policing and Race.” *The Appeal*. January 3, 2018. https://theappeal.org/setting-the-record-straight-on-predictive-policing-and-race-fe588b457ca2/.

“Investigation of the Ferguson Police Department.” 2015. *United States Department of Justice Civil Rights Division*.

Jeremijenko, Natalie. 2005. “Feral Robotic Dogs.” https://www.nyu.edu/projects/xdesign/feralrobots/

Kaste, Martin. 2018. “How Data Analysis Is Driving Policing.” *NPR*. June 25, 2018. https://www.npr.org/2018/06/25/622715984/how-data-analysis-is-driving-policing.

Kaufman, Emily. 2016. “Policing Mobilities through Bio-Spatial Profiling in New York City.” *Political Geography* 55: 72–81. doi:10.1016/j.polgeo.2016.07.006.

Keller, Evelyn Fox. 1995. *Reflections on Gender and Science*. New Haven: Yale University Press.

Kinsley, Samuel. 2014. “The Matter of ‘virtual’ Geographies.” *Progress in Human Geography* 38 (3): 364–84. doi:10.1177/0309132513506270.

Kirsch, Scott. 1995. “The Incredible Shrinking World? Technology and the Production of Space.” *Environment and Planning D: Society and Space* 13 (5): 529–55. doi:10.1068/d130529.

Kitchin, Rob. 2017a. “The Realtimeness of Smart Cities.” *Tecnoscienza: Italian Journal of Science & Technology Studies* 8 (2): 24.

———. 2017b. “Thinking Critically about and Researching Algorithms.” *Information, Communication & Society* 20 (1): 14–29. doi:10.1080/1369118X.2016.1154087.

Kitchin, Rob, and Martin Dodge. 2007. “Rethinking Maps.” *Progress in Human Geography* 31 (3): 331–44. doi:10.1177/0309132507077082.

———. 2011. *Code/Space: Software and Everyday Life*. Software Studies. Cambridge, Mass: MIT Press.

Kitchin, Rob, Justin Gleeson, and Martin Dodge. 2013. "Unfolding Mapping Practices: A New Epistemology for Cartography: *Unfolding Mapping Practices*." *Transactions of the Institute of British Geographers* 38 (3): 480–96. doi:10.1111/j.1475-5661.2012.00540.x.

Kitchin, Rob, Tracey P. Lauriault, and Matthew W. Wilson, eds. 2017. *Understanding Spatial Media*. London: SAGE.

Kittler, Friedrich. 1992. "There Is No Software." *Stanford Literature Review* 9 (1): 81–90.

Klontz, Joshua C., and Anil K. Jain. 2013. "A Case Study of Automated Face Recognition: The Boston Marathon Bombings Suspects." *Computer* 46 (11): 91–94. doi:10.1109/MC.2013.377.

Koper, Christopher S. 1995. "Just Enough Police Presence: Reducing Crime and Disorderly Behavior by Optimizing Patrol Time in Crime Hot Spots." *Justice Quarterly* 12 (4): 649–72. doi:10.1080/07418829500096231.

Kosinski, Michal, David Stillwell, and Thore Graepel. 2013. "Private Traits and Attributes Are Predictable from Digital Records of Human Behavior." *Proceedings of the National Academy of Sciences* 110 (15): 5802–5. doi:10.1073/pnas.1218772110.

Krajeski, Jenna. 2014. "Turkey Vs. Twitter." *The New Yorker*. March 21, 2014. http://www.newyorker.com/news/news-desk/turkey-vs-twitter.

Kuhn, Thomas S. (1962) 2012. *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press.

Kwan, Mei-Po. 2002. "Feminist Visualization: Re-Envisioning GIS as a Method in Feminist Geographic Research." *Annals of the Association of American Geographers* 92 (4): 645–61. doi:10.1111/1467-8306.00309.

———. 2016. "Algorithmic Geographies: Big Data, Algorithmic Uncertainty, and the Production of Geographic Knowledge." *Annals of the American Association of Geographers* 106 (2): 274–82.

Lally, Nick. 2017. "Crowdsourced Surveillance and Networked Data." *Security Dialogue* 48 (1): 63–77. doi:10.1177/0967010616664459.

Lally, Nick, and Ryan Burns. 2017. "Toward a Geographical Software Studies." *Computational Culture*, 6. http://computationalculture.net/special-section-editorial-geographies-of-software/.

Langlois, Ganaele, and Greg Elmer. 2013. "The Research Politics of Social Media Platforms." *Culture Machine* 14: 1–17.

Larkin, Brian. 2013. "The Politics and Poetics of Infrastructure." *Annual Review of Anthropology* 42 (1): 327–43. doi:10.1146/annurev-anthro-092412-155522.

Latour, Bruno. 1986. *Laboratory Life: The Construction of Scientific Facts*. Princeton: Princeton University Press.

———. 2005. *Reassembling the Social an Introduction to Actor-Network-Theory*. Oxford: Oxford University Press.

Law, John. 1992. "Notes on the Theory of the Actor-Network: Ordering, Strategy, and Heterogeneity." *Systems Practice* 5 (4): 379–393.

"Learn About Volkswagen Violations." 2015. Policies and Guidance. *United States Environmental Protection Agency*. October 28, 2015. https://www.epa.gov/vw/learn-about-volkswagen-violations.

Leszczynski, Agnieszka. 2014. "Spatial Media/Tion." *Progress in Human Geography*. doi:10.1177/0309132514558443.

———. 2017. "Digital Methods I: Wicked Tensions." *Progress in Human Geography*. doi:10.1177/0309132517711779.

Leszczynski, Agnieszka, and Sarah Elwood. 2015. "Feminist Geographies of New Spatial Media: Feminist Geographies of New Spatial Media." *The Canadian Geographer / Le Géographe Canadien* 59 (1): 12–28. doi:10.1111/cag.12093.

Levin, Sam. 2015. "Racial Profiling Via Nextdoor.Com." *East Bay Express*. October 7, 2015. https://www.eastbayexpress.com/oakland/racial-profiling-via-nextdoorcom/Content?oid=4526919.

Lorinc, John. 2018. "Busted by Big Data." The Walrus. https://thewalrus.ca/will-big-data-in-crime-fighting-create-a-new-era-of-racial-profiling.

Lum, Kristian, and William Isaac. 2016. "To Predict and Serve?" *Significance* 13 (5): 14–19. doi:10.1111/j.1740-9713.2016.00960.x.

Lyon, David. 2006. "Tear down the walls: on demolishing the panopticon." In *Theorizing surveillance: the panopticon and beyond*, edited by David Lyon. New York: Routledge.

Massanari, Adrienne. 2015. "#Gamergate and The Fappening: How Reddit's Algorithm, Governance, and Culture Support Toxic Technocultures." *New Media & Society*, 1461444815608807.

———. 2015. *Participatory Culture, Community, and Play: Learning from Reddit*. New York: Peter Lang.

Mathew, Ashwin Jacob, and Coye Cheshire. 2012. "The New Cartographers: Trust and Social Order within the Internet Infrastructure." *Proceedings of the 38th Research Conference on Communication, Information and Internet Policy*, January, 2012. https://ssrn.com/abstract=1988216.

Mattern, Shannon. 2015. "History of the Urban Dashboard." *Places Journal*. https://placesjournal.org/article/mission-control-a-history-of-the-urban-dashboard/.

———. 2016. "Cloud and Field- On the Resurgence of 'Field Guides' in a Networked Age." *Places Journal*. https://placesjournal.org/article/cloud-and-field/.

———. 2018. "Databodies in Codespace." *Places Journal*. https://placesjournal.org/article/databodies-in-codespace/.

Medina, Jennifer. 2018. "Website Meant to Connect Neighbors Hears Complaints of Racial Profiling." *The New York Times*, January 19, 2018. https://www.nytimes.com/2016/05/19/us/website-nextdoor-hears-racial-profiling-complaints.html.

Mitchell, Timothy. 1999. "Society, economy, and the state effect." In *State/Culture: State-Formation after the Cultural Turn*, edited by George Steinmetz, 76-97. Ithaca: Cornell University Press.

Mohler, G. O., M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita. 2011. "Self-Exciting Point Process Modeling of Crime." *Journal of the American Statistical Association* 106 (493): 100–108. doi:10.1198/jasa.2011.ap09546.

Mohler, G. O., M. B. Short, Sean Malinowski, Mark Johnson, G. E. Tita, Andrea L. Bertozzi, and P. J. Brantingham. 2015. "Randomized Controlled Field Trials of Predictive Policing." *Journal of the American Statistical Association* 110 (512): 1399–1411. doi:10.1080/01621459.2015.1077710.

Mohler, George O. 2015. Event forecasting system. United States US8949164B1, filed September 6, 2012, and issued February 3, 2015. https://patents.google.com/patent/US8949164B1/en.

Monahan, Torin, and Jennifer T. Mokos. 2013. "Crowdsourcing Urban Surveillance: The Development of Homeland Security Markets for Environmental Sensor Networks." *Geoforum* 49 (October): 279–88. doi:10.1016/j.geoforum.2013.02.001.

Monk, Janice, and Susan Hanson. 1982. "ON NOT EXCLUDING HALF OF THE HUMAN IN HUMAN GEOGRAPHY." *The Professional Geographer* 34 (1): 11–23. doi:10.1111/j.0033-0124.1982.00011.x.

Montgomery, David, Sari Horwitz, and Marc Fisher. 2013. "Police, Citizens and Technology Factor into Boston Bombing Probe." *The Washington Post*, April 20, 2013. 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.

Montjoye, Yves-Alexandre de, César A. Hidalgo, Michel Verleysen, and Vincent D. Blondel. 2013. "Unique in the Crowd: The Privacy Bounds of Human Mobility." *Scientific Reports* 3 (1). doi:10.1038/srep01376.

Muchnik, Lev, Sinan Aral, and Sean J. Taylor. 2013. "Social Influence Bias: A Randomized Experiment." *Science* 341 (6146): 647–51. doi:10.1126/science.1240466.

Neman, Lily Hay. 2016. "What We Know About Friday's Massive East Coast Internet Outage." *WIRED*. October 21, 2016. https://www.wired.com/2016/10/internet-outage-ddos-dns-dyn.

Ney, Peter, Ian Smith, Gabriel Cadamuro, and Tadayoshi Kohno. 2017. "SeaGlass: Enabling City-Wide IMSI-Catcher Detection." *Proceedings on Privacy Enhancing Technologies* 2017 (3). doi:10.1515/popets-2017-0027.

Nost, Eric. 2015. "Performing Nature's Value: Software and the Making of Oregon's Ecosystem Services Markets." *Environment and Planning A* 47 (12): 2573–90. doi:10.1177/0308518X15616631.

O'Donovan, Caroline. 2017. "Racial Profiling Is Still A Problem On Nextdoor." *BuzzFeed*. May 18, 2017. https://www.buzzfeed.com/carolineodonovan/racial-profiling-is-still-a-problem-on-nextdoor.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

O'Sullivan, David. 2018. "Big Data: Why (Oh Why?) Computational Social Science?" In *Thinking Big Data in Geography: New Regimes, New Research*, edited by Jim Thatcher, Josef Eckert, and Andrew Shears. Lincoln, NE: University of Nebraska Press.

O'Sullivan, David, Luke Bergmann, and Jim E. Thatcher. 2017. "Spatiality, Maps, and Mathematics in Critical Human Geography: Toward a Repetition with Difference." *The Professional Geographer*, June, 1–11. doi:10.1080/00330124.2017.1326081.

Payne, Will. 2017. "Welcome to the Polygon: Contested Digital Neighborhoods and Spatialized Segregation on Nextdoor." *Computational Culture* 6. http://computationalculture.net/welcome-to-the-polygon-contested-digital-neighborhoods-and-spatialized-segregation-on-nextdoor/.

Perry, Walter, Brian McInnis, Carter Price, Susan Smith, and John Hollywood. 2013. *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. RAND Corporation. doi:10.7249/RR233.

Peters, John Durham. 2015. *The Marvelous Clouds: Toward a Philosophy of Elemental Media*. Chicago: University of Chicago Press.

Picon, Antoine. 2003. "Nineteenth-Century Urban Cartography and the Scientific Ideal: The Case of Paris." *Osiris* 18 (January): 135–49. doi:10.1086/649381.

Pinch, Trevor J., and Wiebe E. Bijker. 1987. "The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other." In *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, edited by Trevor J. Pinch and Wiebe E. Bijker.

Poiker, Tom. 1995. "Preface." *Cartography and Geographic Information Systems* 22 (1): 3–4. doi:10.1559/152304095782540591.

Polvi, Natalie, Terah Looman, Charlie Humphries, and Ken Pease. 1991. "THE TIME COURSE OF REPEAT BURGLARY VICTIMIZATION." *The British Journal of Criminology* 31 (4): 411–14. doi:10.1093/oxfordjournals.bjc.a048138.

Privacy International. 2017. "Fintech: Privacy and Identity in the New Data-Intensive Financial Sector." https://privacyinternational.org/sites/default/files/2017-12/Fintech%20report.pdf.

Rancière, Jacques. 2006. *The Politics of Aesthetics: The Distribution of the Sensible*. London; New York: Continuum.

Ratcliffe, Jerry H. 2015. "Harm-Focused Policing." *Ideas in American Policing* 19.

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

Rehnquist. 2000. Illinois v. Wardlow (Opinion of the Court), 528 U.S. 119. U.S. Supreme Court.

Robinson, David, and Logan Koepke. 2016. "Stuck in a Pattern: Early Evidence on 'Predictive Policing' and Civil Rights." *Upturn*. August, 2016. https://www.teamupturn.org/static/reports/2016/stuck-in-a-pattern/files/Upturn_-_Stuck_In_a_Pattern_v.1.01.pdf.

Rose, Gillian. 2016. "Rethinking the Geographies of Cultural 'objects through Digital Technologies: Interface, Network and Friction." *Progress in Human Geography* 40 (3): 334–51. doi:10.1177/0309132515580493.

Rose-Redwood, Reuben. 2012. "With Numbers in Place: Security, Territory, and the Production of Calculable Space." *Annals of the Association of American Geographers* 102 (2): 295–319. doi:10.1080/00045608.2011.620503.

Roth, Robert E. 2013. "Interactive Maps: What We Know and What We Need to Know." *Journal of Spatial Information Science*, no. 6 (June). doi:10.5311/JOSIS.2013.6.105.

Salter, M. B. 2006. "The Global Visa Regime and the Political Technologies of the International Self: Borders, Bodies, Biopolitics." *Alternatives: Global, Local, Political* 31 (2): 167–89. doi:10.1177/030437540603100203.

Saunders, Jessica, Priscillia Hunt, and John S. Hollywood. 2016. "Predictions Put into Practice: A Quasi-Experimental Evaluation of Chicago's Predictive Policing Pilot." *Journal of Experimental Criminology* 12 (3): 347–71. doi:10.1007/s11292-016-9272-0.

Schneider, Christopher J., and Daniel Trottier. 2012. "The 2011 Vancouver Riot and the Role of Facebook in Crowd-Sourced Policing." *BC Studies* 175.

Schuurman, Nadine. 2000. "Trouble in the Heartland: GIS and Its Critics in the 1990s." *Progress in Human Geography* 24 (4): 569–90. doi:10.1191/030913200100189111.

Schuurman, Nadine, and Geraldine Pratt. 2002. "Care of the Subject: Feminism and Critiques of GIS." *Gender, Place and Culture: A Journal of Feminist Geography* 9 (3): 291–299.

Shapiro, Aaron. 2017. "Reform Predictive Policing." *Nature* 541 (7638): 458–60. doi:10.1038/541458a.

———. 2018. "Design, Control, Predict: Cultural Politics in the Actually Existing Smart City." PhD diss., University of Pennsylvania.

Shelton, Taylor, Ate Poorthuis, and Matthew Zook. 2015. "Social Media and the City: Rethinking Urban Socio-Spatial Inequality Using User-Generated Geographic Information." *Landscape and Urban Planning* 142: 198–211. doi:10.1016/j.landurbplan.2015.02.020.

Shelton, Taylor, Matthew Zook, and Alan Wiig. 2014. "The 'Actually Existing Smart City.'" *Cambridge Journal of Regions, Economy and Society*.

Sherman, Lawrence W. 1998. "Evidenced-Based Policing." *Ideas in American Policing*.

———. 2015. "A Tipping Point for 'Totally Evidenced Policing': Ten Ideas for Building an Evidence-Based Police Agency." *International Criminal Justice Review* 25 (1): 11–29. doi:10.1177/1057567715574372.

Sherman, Lawrence W., and David Weisburd. 1995. "General Deterrent Effects of Police Patrol in Crime 'Hot Spots': A Randomized, Controlled Trial." *Justice Quarterly* 12 (4): 625–48. doi:10.1080/07418829500096221.

Somers, James. 2017. "The Coming Software Apocalypse." *The Atlantic*. September 26, 2017. https://www.theatlantic.com/technology/archive/2017/09/saving-the-world-from-code/540393/.

Star, Susan Leigh. 1999. "The Ethnography of Infrastructure." *American Behavioral Scientist* 43 (3): 377–91. doi:10.1177/00027649921955326.

Star, Susan Leigh, and Ruhleder, Karen. 1996. "Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces." *Information Systems Research* 7 (1): 111–34.

Steyerl, Hito. 2014. "Proxy Politics: Signal and Noise." *E-Flux Journal* 60. http://www.e-flux.com/journal/60/61045/proxy-politics-signal-and-noise/.

Striphas, Ted. 2015. "Algorithmic Culture." *European Journal of Cultural Studies* 18 (4–5): 395–412.

Suchman, Lucy A. 1985. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Palo Alto, CA: Xerox Palo Alto Research Center.

Surowiecki, James. 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.

———. 2013. "The Wise Way to Crowdsource a Manhunt." *The New Yorker.* April 23, 2013. 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.* December 15, 2015. https://www.washingtonpost.com/news/the-fix/wp/2015/12/15/who-said-what-and-what-it-meant-the-fifth-gop-debate-annotated/

Tewksbury, Doug. 2012. "Crowdsourcing homeland security: The Texas virtual borderwatch and participatory citizenship." *Surveillance and Society, 10*(3-4), 249-262. doi: 10.24908/ss.v10i3/4.3464

Thatcher, Jim, David O'Sullivan, and Dillon Mahmoudi. 2016. "Data Colonialism through Accumulation by Dispossession: New Metaphors for Daily Data." *Environment and Planning D: Society and Space* 34 (6): 990–1006. doi:10.1177/0263775816633195.

Thatcher, Jim, Luke Bergmann, Britta Ricker, Reuben Rose-Redwood, David O'Sullivan, Trevor J. Barnes, Luke R. Barnesmoore, et al. 2016. "Revisiting Critical GIS." *Environment and Planning A* 48 (5): 815–24. doi:10.1177/0308518X15622208.

"The Philadelphia Predictive Policing Experiment." 2018. Temple University. http://www.cla.temple.edu/cj/center-for-security-and-crime-science/the-philadelphia-predictive-policing-experiment.

Thornton, Pip. 2017. "Geographies of (Con)Text: Language and Structure in a Digital Age." *Computational Culture* 6. http://computationalculture.net/geographies-of-context-language-and-structure-in-a-digital-age.

Tobin, Ariana. 2017. "Why We Had to Buy Racist, Sexist, Xenophobic, Ableist." *ProPublica*. November 27, 2017. https://www.propublica.org/article/why-we-had-to-buy-racist-sexist-xenophobic-ableist-and-otherwise-awful-facebook-ads.

Tobler, Waldo Rudolph. 1961. "Map Transformations of Geographic Space." PhD diss., University of Washington

Trottier, Daniel. 2014. "Crowdsourcing CCTV Surveillance on the Internet." *Information, Communication & Society* 17 (5): 609–26. doi:10.1080/1369118X.2013.808359.

Tuan, Yi-Fu. 1976. "Humanistic Geography." *Annals of the Association of American Geographers* 66 (2): 266–76.

United States Department of Homeland Security n.d. "If You See Something, Say Something." http://www.dhs.gov/see-something-say-something.

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. http://www.gao.gov/assets/660/658923.pdf.

Van Dijck, José. 2009. "Users like You? Theorizing Agency in User-Generated Content." *Media, Culture, and Society* 31 (1): 41.

Vincent, Kristen, Robert E Roth, Sarah A Moore, Qunying Huang, Nick Lally, Carl M Sack, Eric Nost, and Heather Rosenfeld. 2018. "Improving Spatial Decision Making Using Interactive Maps: An Empirical Study on Interface Complexity and Decision Complexity in the North American Hazardous Waste Trade." *Environment and Planning B*, 18.

Wardrip-Fruin, Noah. 2012. *Expressive Processing: Digital Fictions, Computer Games, and Software Studies*. Cambridge, MA: MIT Press.

Wei, Wang. 2017. "New Rapidly-Growing IoT Botnet Threatens to Take Down the Internet." *The Hacker News*. October 20, 2017. http://thehackernews.com/2017/10/iot-botnet-malware-attack.html.

Weisburd, David, and Cody W. Telep. 2014. "Hot Spots Policing: What We Know and What We Need to Know." *Journal of Contemporary Criminal Justice* 30 (2): 200–220. doi:10.1177/1043986214525083.

Weisburd, David, Laura A. Wyckoff, Justin Ready, John E. Eck, Joshua C. Hinkle, and Frank Gajewski. 2006. "DOES CRIME JUST MOVE AROUND THE CORNER? A CONTROLLED STUDY OF SPATIAL DISPLACEMENT AND DIFFUSION OF CRIME CONTROL BENEFITS." *Criminology* 44 (3): 549–92. doi:10.1111/j.1745-9125.2006.00057.x.

Weiser, Mark. 1991. "The Computer for the 21st Century." *Scientific American* 265 (3): 94–104. doi:10.1038/scientificamerican0991-94.

Woods, Louis Lee. 2012. "The Federal Home Loan Bank Board, Redlining, and the National Proliferation of Racial Lending Discrimination, 1921–1950." *Journal of Urban History* 38 (6): 1036–59. doi:10.1177/0096144211435126.

Woodward, Keith. 2016. "State Reason and State Affect." *Political Geography* 51 (March): 89–91. doi:10.1016/j.polgeo.2016.01.005.

Woodward, Keith, John Paul Jones III, and Sallie A Marston. 2010. "Of Eagles and Flies: Orientations toward the Site: Of Eagles and Flies: Orientations toward the Site." *Area*, January, no-no. doi:10.1111/j.1475-4762.2009.00922.x.

Woolf, Nicky. 2016. "DDoS Attack That Disrupted Internet Was Largest of Its Kind in History, Experts Say." *The Guardian*. October 26, 2016. https://www.theguardian.com/technology/2016/oct/26/ddos-attack-dyn-mirai-botnet.

Woolgar, Steve. 1990. “Configuring the User: The Case of Usability Trials.” *The Sociological Review* 38 (S1): 58–99.

Wu, Hao-Yu, Michael Rubinstein, Eugene Shih, John Guttag, Frédo Durand, and William Freeman. 2012. “Eulerian Video Magnification for Revealing Subtle Changes in the World.” *ACM Transactions on Graphics* 31 (4): 1–8. doi:10.1145/2185520.2185561.

Wu, Yuning, Ivan Y. Sun, and Ruth A. Triplett. 2009. “Race, Class or Neighborhood Context: Which Matters More in Measuring Satisfaction with Police?” *Justice Quarterly* 26 (1): 125–56. doi:10.1080/07418820802119950.

York, Jillian C. 2014. “Why Is Turkey Blocking Twitter?” *Electronic Frontier Foundation*. March 20, 2014. https://www.eff.org/deeplinks/2014/03/why-turkey-blocking-twitter.

Young, Stephen, Alasdair Pinkerton, and Klaus Dodds. 2014. “The Word on the Street: Rumor, ‘Race’ and the Anticipation of Urban Unrest.” *Political Geography* 38 (January): 57–67. doi:10.1016/j.polgeo.2013.11.001.

Zook, Matthew, Mark Graham, Taylor Shelton, and Sean Gorman. 2010. “Volunteered Geographic Information and Crowdsourcing Disaster Relief: A Case Study of the Haitian Earthquake.” *World Medical & Health Policy* 2 (2): 6–32. doi:10.2202/1948-4682.1069.