DATA-DRIVEN ACCOUNTABILITY: EXAMINING AND REORIENTING THE MYTHOLOGIES OF DATA

Nitya Verma

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the School of Informatics and Computing,
Indiana University

May 2020

Accepted by the Graduate Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

____
Lynn Dombrowski, Ph.D., Chair

____
Davide Bolchini, Ph.D.

April 25, 2019

____
Alyson Young, Ph.D.

____
Peter Seybold, Ph.D.

____
Amy Voida, Ph.D.

____
Michael Muller, Ph.D.

DEDICATION

To my family, friends, and mentors.

ACKNOWLEDGEMENT

I would first like to thank my parents for their motivation and support for my graduate study. This manuscript would not be possible if they did not continuously hold space and show up for me during times of stress. I would also like to thank my sister for her support. I am grateful for my friends both in and out of IUPUI for wishing me well and being a vital part of my support system.

I am grateful for the opportunity to pursue graduate school in the first place, which has also ended up being one of my most cherished endeavors yet. For that, I would like to thank my advisor, Lynn Dombrowski, for her unwavering guidance and perseverance in my doctoral training. She continues to push me to be more robust in research, always setting me up for higher standards, both professionally and personally. I appreciate her allowing me the freedom to pursue this research topic-- always meeting me where I am, but pushing me to produce better quality of work. She has invested countless hours into training me to be a more skilled and thoughtful researcher. Enumerating Lynn’s role in my doctorate journey is part of a longer narrative, and this acknowledgement by no means measures up. I am deeply grateful for her taking me under her wing and for her confidence in me. Thank you for everything, always.

I would also like to thank Davide Bolchini as my co-advisor. His passion and enthusiasm for HCI and his students is infectious and inspiring. He has always provided multiple ways of support within the department to help me progress in this doctorate. One of the best things about him as a teacher is how remarkably he listens, understands, and creates a more relevant and beneficial experience for his students. I have learned a lot from him about staying grounded in the principles of HCI as an interdisciplinary field. Thank you for your positivity and guidance.

Sincere thanks to my committee member and first advisor, Amy Voida. I consider Amy to be touchstone of excellence in HCI and am immensely grateful to have started off my PhD under her guidance. My time as her advisee set me up with tools I continue to come back to as I navigated research work and academia, in general. Thank you, Amy.

A most sincere thank you to my committee members: Alyson Young, Peter Seybold, and Michael Muller. I deeply appreciate your intellectual investment in this research project and me, as a scholar. I also want to take a moment to thank all my teachers and mentors that came before this doctorate, who encouraged and pushed me to expand my boundaries.

Most sincere thanks to my friends, LaRahia and Anisha, for being such consistent pillars of support through most of my life. Thank you for being there for me. Also, thanks to Milka, Oindrilla, and Vidhur for understanding and going through this doctoral process with me. I would like to thank Dylan for encouraging and supporting me in taking on this challenge five years ago.

Finally, I am immensely grateful for Roland and his slobbery companionship.

Nitya Verma

DATA-DRIVEN ACCOUNTABILITY: EXAMINING AND REORIENTING THE MYTHOLOGIES OF DATA

In this work, I examine and design sociotechnical interventions for addressing limitations around data-driven accountability, particularly focusing on politically contentious and systemic social issues (*i.e.,* police accountability). While organizations across sectors of society are scrambling to adopt data-driven technologies and practices, there are epistemological and ethical concerns around how data use influences decision-making and actionability. My work explores how stakeholders adopt and handle the challenges around being data-driven, advocating for ways HCI can mitigate such challenges.

In this dissertation, I highlight three case studies that focus on data-driven, human-services organizations, which work with at-risk and marginalized populations. First, I examine the tools and practices of nonprofit workers and how they experience the mythologies associated with data use in their work. Second, I investigate how police officers are adopting data-driven technologies and practices, which highlights the challenges police contend with in addressing social criticisms around police accountability and marginalization. Finally, I conducted a case study with multiple stakeholders around police accountability to understand how systemic biases and politically charged spaces perceive and utilize data, as well as to develop the design space around how alternative futures of being data-driven could support more robust and inclusive accountability. I examine how participants situate the concepts of power, bias, and truth in the data-driven practices and technologies used by and around the police.

With this empirical work, I present insights that inform the HCI community at the intersection of data design, practice, and policies in addressing systemic social issues.

Lynn Dombrowski, Ph.D., Chair

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI: Artificial Intelligence

BI: Business Intelligence

HCI: Human-Computer Interaction

MMPD: Midwest Metropolitan Police Department

POC: People of Color

UCR: Uniform Crime Reporting

P#: Participant Number

Chapter 1: Introduction

Human-Computer Interaction (HCI) focuses on the intersection of humans and technology, with increasing attention being paid to how technology impacts larger, societal issues (*e.g.,* social justice, health and wellness, criminal justice, etc.) [Baumer & Silberman, 2011]. Similarly, critics of big data have indicated that data-driven technologies in particular (*i.e.,* business intelligence tools, predictive tools, AI decision-making support tools, etc.) stand to exacerbate existing biases, which is disproportionately consequential for at-risk and marginalized populations around social issues of poverty, criminal justice, mental health, etc. [boyd & Crawford, 2013; Eubanks, 2018]. However, because quantitative data enables organizations to rationalize efforts and legitimize decision-making [Morgan, 1997], public and private sector organizations[1] have been scrambling to adopt data-driven practices and technologies [Williams, 2015; Delgado & Stefancic, 2015]. For human-services agencies, particularly like the police, who are under intense criticisms, being data-driven can provide ways to better serve at-risk people and increase accountability. In the context of this research, I define accountability as a condition of being responsible and answerable to the public for *what* and *how* workers perform on an agency and individual level, including how they do and do not equitably work towards their organizational mission, with respect to the law [Walker, 2001].

Subsequently, as data-driven practices becomes more prevalent and promises improved accountability, researchers voice concerns about the data use limitations as it stands to change how we define, create, and engage with knowledge, particularly in

[1] Global revenue from BI and analytics tools was projected to reach $16.9 billion in 2016, a 5.2% annual increase [25].

human-services contexts, where people's unpredictable nature is inherent [boyd & Crawford, 2012; Marshall et. al., 2016; Verma & Voida, 2016]. Researchers, for example, argue that the shift towards quantitative data changes assumptions about the meaning of knowledge and about how people "should" engage with information [boyd & Crawford, 2012]. They contend that there is an increasingly pervasive "mythology" of big data:

> …the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy. [boyd & Crawford, 2012]

This mythology must become more transparent in research about data-driven technology since these systems are the most prominent user-facing manifestation of 'big data' and its related computational turn in thinking within organizations. Understanding these mythologies and how they manifest on-the-ground for stakeholders is necessary in order to mitigate the negative impacts of being data-driven, especially in politically contentious spaces (*i.e.,* crime, poverty, mental health, etc.). My dissertation addresses the gap in the literature of studying the human experience and impact of technologies that manifest the computational turn. In doing so, I contribute at the intersection of design, practice, and policies for being data driven [Jackson et. al. 2014]. Through this research, I have identified the mythologies of being data-driven and how data use can constrain the space for action within an organization; the limitations around how data politics inhibit trust from stakeholders; and lastly, how stakeholders experience data use as an extension of police's power in the quest for accountability. Subsequently, this dissertation contributes design guidelines and implications for designers, practitioners, and policy-makers to help mitigate the limitations and biases associated with data-driven technologies and practices.

More specifically, I use data-driven human-services organizations as a case site for investigating the limitations and challenges around data use. I use two cases of human-services organizations—nonprofit and police—to understand how the participants experienced being data-driven as organizational values [Verma & Voida, 2016], as well as the challenges and limitations they faced in adopting data-driven technologies and practices [Verma & Dombrowski, 2018]. Particularly, most of this dissertation investigates the problem and design space around police's data use. Data-driven policing promises to improve police accountability through eliminating "undesirable biases," introducing "more fairness into the police decision-making process," and improving internal and external accountability practices [Davis, et. al., 2016; Joh, 2014]. However, while data-driven policing can improve accountability in some ways, there are limitations in how such data practices create the conditions for and ensure accountability to the public, including how data can be used to shape and legitimize agendas, frame narratives, and perpetuate existing biases and inequalities [Verma & Dombrowski, 2018].

These design challenges, identified in my research, are even more important to address when considering the impact data use has for at-risk people and marginalized communities around historically and politically contentious social issues of crime, drug abuse, mental illness, and poverty. As evident from previous research, technology and data are not neutral nor free from values, interests, and biases. No data is truly "raw"; the identification and decisions of what data is to be measured and how data is categorized are political acts, motivated implicitly or explicitly by different values [Crawford, 2013;

Ribes & Jackson, 2013]. Values and biases are embodied through the design of the systems and practices as data is produced and used by police [LeDantec et. al., 2009; Voida et. al., 2014; Friedman et. al., 2006; Swenson, 2014].

While politics and systemic inequalities are responsible for creating social issues, such as police accountability, I contend that technology use can and must play an important role in dismantling existing biases and inequalities instead of exacerbating and perpetuating them. With this research, I advocate ways for the biases and agendas embedded in technology to be evaluated and mitigated as data-driven technology and practices become more prevalent in our society. In order to address these limitations, this dissertation aims to reorient the mythologies around data use (*e.g.,* objectivity, transparency, and trust) into pragmatic design guidelines for the HCI community to consider. For example, while using data in an objective and neutral manner is implausible, there are sociotechnical interventions that can strengthen the objectivity in data representations, by being inclusive of key perspectives around data use. I discuss design implications for how the concept of 'strong objectivity' can be incorporated into data-driven technologies and practices [Harding, 1995] to help mitigate the power disparities that emerge with a single group of people shaping data (and subsequently, spaces for action). I contribute guidelines for achieving strong objectivity at the intersection of data design, practice and policies. Similarly, I also present design implications for what it means for HCI to discuss politically contentious social issues that revolve around systemic power disparities, especially since being data-driven is adopted with the goal of improving trust. In such cases, I argue that using data for police accountability with an attitude of *mistrust* is more conducive because there are historical and systemic reasons for the absence of trust between police and POC communities. Mistrust is an appropriate value in certain contexts [Carey, 2017], and I argue an important one for addressing and mitigating the existing biases in police accountability narratives. With mistrust as a value in design, I advocate that data representations can more satisfactorily address the social scrutiny police face, for instance. Subsequently and finally, I argue that while data can be used to hold organizations and individuals accountable, there are design opportunities to hold data use accountable as well. For that end, I present design guidelines for information transparency, which here refers to disclosing constraints and criterion around data production and usage. Information transparency is an important key in creating more robust and inclusive accountability between organizations and their stakeholders.

Ultimately, my dissertation addresses the following research questions around the limitations of data-driven technologies and practices in human-services contexts:

1. How do the mythologies of data use manifest as organizational values for stakeholders? How does data stand to rationalize and legitimize spaces for action for these stakeholders?
2. What are the challenges stakeholders experience in using data-driven technologies for shaping accountability?
3. What are key value tensions for accountability between groups of people with large power disparities (*i.e.,* police and marginalized communities)?

1.1 Dissertation Outline

Chapter 2: Literature Review

This research draws from literature in five areas of related work including information management in human-services organizations (nonprofit & police), data-driven policing, data politics, community informatics, and accountability in government and HCI.

Chapter 3: Methods

In this three phase research, I conducted case studies mainly using qualitative methods to better understand data-driven technologies and practices, related sociotechnical challenges, and possible design interventions for human-services. This dissertation is a product of these research activities. By utilizing these qualitative methods, I am able to identify the practical and on-the-ground limitations and challenges data use and how data-driven rhetoric is intertwined with larger principles of truth, accountability, and power.

Chapter 4: Mythologies of Data-Driven Technologies and Practices

In this chapter, I present results from my first case study of the use of data-driven systems in a nonprofit, human services organization. I characterize four mythologies of data-driven systems that participants experience as shared organizational values and are core to their trajectory towards a "culture of data": data-driven, predictive and proactive, shared accountability, and inquisitive. For each mythology, I also discuss the ways in which being actionable is impeded by a disconnect between the aggregate views of data that allows them to identify areas of focus for decision making and the desired "drill down" views of data that would allow them to understand how to act in a data-driven context. These findings contribute initial empirical evidence for the impact of data-driven technology's epistemological biases on organizations and suggest implications for the design of technologies to better support data-driven decision making by legitimizing non-traditional forms of data.

Chapter 5: Challenges and Limitations in Data-Driven Policing

In this chapter, I present results from my second qualitative field study about the adoption of data-driven policing strategies in a Midwestern police department in the United States. Proponents of data-driven policing strategies claim that it makes policing organizations more effective, efficient, and accountable and has the potential to address some policing social criticisms (*e.g.* racial bias, lack of accountability and training). What remains less understood are the challenges when adopting data-driven policing as a response to these criticisms. Here, I identify three key challenges police face with data-driven adoption efforts: *data-driven frictions*, *precarious and inactionable insights*, and *police metis concerns*. I demonstrate the issues that data-driven initiatives create for policing and the open questions police agents face. These findings contribute an empirical account of how policing agents attend to the strengths and limits of big data's knowledge claims, as well as help me develop the problem space for how biases manifest in data-driven policing, as well as how data-driven policing can hinder accountability.

Chapter 6: Re-Orienting Data-Driven Mythologies

In this chapter, I present results from the final qualitative field study with various stakeholders around police accountability, including police officers, anti-police activists, and various community members. While data-driven policing is meant to improve police accountability and their relationships with the communities, my results demonstrate that stakeholders' concerns around data use for police accountability revolve around how data stands to be an extension of police's power. These key issues emerge around who has the power to create and legitimize data as knowledge, who gets to utilize that data to receive and allocate resources and services, and finally, how police's data use stands to exacerbate the policing and surveillance of marginalized communities. These empirical findings contribute to how the limitations around data bias stand to impede accountability efforts, as well as the design space for mitigating those limitations.

Chapter 7: Design Implications

In this section, I pull together the empirical work to demonstrate the design space around data-driven technology and practices. Ultimately, this research contributes to the calls for the HCI community to consider design implications around the limitations of data use and what it means for the human-beings being served by such organizations. Here, I contend that biases and agendas embedded in technology need to be evaluated and mitigated as data-driven technology and practices become more prevalent in our society. In order to address these limitations, I aim to reorient the mythologies around data use (objectivity, transparency, and trust) into pragmatic design guidelines for the HCI community to mitigate the inherent biases in data use. To mitigate the mythology of objectivity, I use the concept of strong objectivity to demonstrate how designers and practitioners could mitigate the power disparities in whose voices are represented in the design, practice, and policies around data. To address the issue of data framing narratives, I advocate for designing with principles of information transparency. Finally, I argue that to improve data's credibility further, particularly in contentious and power disparate contexts, mistrust is a more appropriate value to adopt to guide data practices. Ultimately, when bias can be understood as situated in a larger system, rather than an individual shortcoming, we as a society can become better equipped to mitigate the disparities that emerge through bias. This dissertation aims to provide sociotechnical interventions for the HCI community at the intersection of data design, practices, and policies.

Chapter 8: Conclusion

Through these three case studies, I use human-services organizations' adoption of data-driven technologies to examine the limitations of data use for accountability and mitigating power disparities. This chapter concludes with broader implications around the sociotechnical interventions for being data-driven in more robust ways, as well as future directions for research.

Chapter 2: Background and Literature Review

This research draws from the following areas of related work: information management in human-services organizations (nonprofit & police); more specifically, data-driven policing; data politics; accountability in government and HCI. These literature sections will help situate the following research questions around the limitations of data-driven technologies and practices in human-services contexts:

1. How do the mythologies of data use manifest as organizational values for stakeholders? How does data stand to rationalize and legitimize spaces for action for these stakeholders?
2. What are the challenges stakeholders experience in using data-driven technologies for shaping accountability?
3. What are key value tensions for accountability between groups of people with large power disparities (*i.e.,* police and marginalized communities)?

2.1 Information Management in Human-Services

This dissertation explores data use in human-services through two kinds of case sites: nonprofit and police organizations. While the missions and the influence of these organizations varied, both organizations work with at-risk people, with similar sociotechnical constraints (i.e., funding, man-power, IT infrastructure, etc.). In this section, I first briefly outline literature on data and information management in the nonprofit sector, then discuss the background and role of technology in police.

Nonprofit Sector

The nonprofit sector serves many critical functions and offers services that are underprovided by the government and the for-profit sector [Mair et. al., 2015; Salamon &

Sokolowski, 2004; Burkholder, 1992; Miller et. al., 2007; Reeves & Burt, 2006]. Nonprofit organizations are under increasing pressure to demonstrate their performance and impact to funding agencies (e.g. [Yin, 1994]). So while data collection is a substantial part of the work that most nonprofits do, there is increasing evidence of the costs: "Nonprofits are often collecting heaps of dubious data, at great cost to themselves and ultimately to the people they serve" [Snibbe, 2006]. Research about performance and accountability in nonprofit organizations suggests that as data collection becomes the focus, data-collection tools can diminish the quality of service to clients leading to less effective performance [Benjamin & Campbell, 2014; Kong, 2008; Benjamin, 2008]. Kong also notes that it is not helpful to apply management strategies that work in the for-profit organization to organizations in the nonprofit sector because those strategies typically fail to address the social dimension of mission-driven organizations [Kong, 2008].

More generally, nonprofits often operate under significant constraints in technical resources and expertise that can make collecting, managing, and using data a challenging endeavor [Voida et. al., 2012; Herman, 1994].

Police

While it is beyond the scope of this dissertation to discuss policing history in depth, here, I focus on the evolving policing philosophies related to law enforcement's purpose in the US. Further, I situate how both changing technological landscape and policing philosophies impact technology choices and outcomes in this section. Manning describes the role police in America as a bureaucratic organization that enforces political order through force [Manning, 2005]. However, police's legitimacy is not infallible, their practices have been subject to criticisms and called to change. Since the 1960s, citizens have made calls to reform police efforts based on the traditional policing model [Archbold, 2012; Sparrow, 1988]. Traditional policing refers to "responding to calls for service and managing crimes in a reactive manner." The traditional policing model values quantitative factors like arrest numbers and police officer response time, but not necessarily issues related to a community's quality of life [Archbold, 2012]. Many of these call for reform want police efforts to focus on a community policing approach, which is distinguished by collaborations between police and community; management, personnel, and information technology support; and lastly, a problem-solving orientation to develop and evaluate police strategies to reduce crime instead of solely enforcing the law [Archbold, 2012; Cops, 2011]. According to the Bureau of Justice Statistics, about 70% of police departments in 2013 included at least one community policing component in their mission, an increase from 53% in 2007 [Reaves, 2013]. Community policing refers to the "systematic use of partnerships and problem-solving techniques between the police and the community. These strategies proactively address the immediate conditions that give rise to public safety issues such as crime, social disorder, and fear of crime." These calls for reform coincide with a diminishing perception of police legitimacy and police-community relationships in the US [Kochel, 2011; Lind & Tyler, 1988; Sunshine & Tyler, 2003; Tyler, 2006].

While researchers suggest that information technology (IT) adoption offers benefits like better administrative support, improved productivity and increased efficiency, scholars have raised questions about how police work computerization impacts community outcomes [Nunn, 2001; Chan, 2001; Lewis & Lewis, 2011]. These concerns focus on detrimental changes to policing outcomes and efforts associated with computerization. The first major concern is that IT adoption by itself does not increase productivity for police unless it is paired with strategic management practices situated in addressing community-specific problems (*i.e.,* evidence-driven, decision-making tools) [Garicano & Heaton, 2010]. Secondly, while technology is meant to save time, it appears that in efforts to be more efficient, data work responsibilities shift to undertrained officers. For example, while laptops in patrol cars have provided on-the-ground, timely access to information[2], the responsibility for data collection has shifted from over-the-phone transcribers to the on-the-ground officers, which raises concerns about data accuracy and quality and the lack of technical training [Northrop, et. al., 1995]. More generally, law enforcement, like other public services, operates under substantial technical, expertise, and budgetary constraints [Nunn, 2001]. This research contributes to this space an understanding of how police stakeholders are thinking through and dealing with the social criticisms from the public, while focusing on how data and technology influence their internal and external accountability practices.

2.2 Data-Driven Policing

To fully understand how this research looks at police specifically as a case site for being data-driven, in this section, I present literature on the emergence of data-driven policing and its advantages and challenges. Data-driven policing refers to police decision making processes guided by evidence-based strategies to understand crime problems and allocate limited resources with the goals of greater effectiveness and efficiency in order to reduce criminal and social harm [Hardy, 2010]. While definitions vary, researchers

[2]According to a 2015 report by the Bureau of Justice Statistics, 9 in 10 police officers in the US are provided with in-field computerized access to police records [6].

indicate how data-driven approaches relate to and embedded in community-oriented policing [Peak & Barthe, 2009]. Here, I use 'data-driven policing' to refer to the multifaceted, sociotechnical practices situated within the tool ecology and organizational practices for data-driven decision making.

Data-driven policing is characterized by having the following components: data collection, analysis, local partnerships with community stakeholders, strategic operations, information sharing, monitoring, evaluating and adjusting operations, and measuring outcomes [Hardy, 2010]. Due to the prevalence of open forms of relevant data (*e.g.,* open governmental data; social media), data collection and analysis is not constrained to only the data collected by police, but police also frequently reuse other public data (*i.e.* census, fire department, medical services). Big data in policing refers to the amalgamation of data sources including camera feeds, license plate readers, alert-based systems, police reports, public services data, crime tips, and social media. Researchers characterize big data's role in policing to be predictive (applying a computational model to historical data to determine criminal activity's future likelihood), repurpose already available data, and engage in mass surveillance [Joh, 2014]. While police agencies have conducted crime analysis on their own historical data to track trends, data-driven policing is distinguished by the characteristics related to big data (*i.e.,* large volumes, variety of data at much more real-time velocities). Data-driven policing is also, in part, the police's response to social criticisms of police practices in attempts to be more proactive and accountable [Manning, 2001]. However, researchers have called for more research on how big data can support police accountability [Brayne, 2017]. I build on this work to demonstrate the challenges in using data-driven strategies to address criticism through designing from police and community perspectives.

Law enforcement research identifies many benefits of data-driven policing including optimizing police resources, eliminating human biases, encouraging accountability within the police agencies and to the community, problem-solving, improving community support and engagement, and reducing crime [Kochel, 2011; Dabney, 2010; Brayne, 2017; Weisburd, 2003]. Researchers argue that by implementing data-driven policing, technology eliminates "undesirable biases" to "introduce more fairness into the police decision-making process" as a means to address recent social criticisms against police [Joh, 2014]. While organizational changes like committed management, defined objectives, and data expertise can help organizations work towards these benefits, researchers have called for further research to understand *how* data-driven policing strategies impact police practices, stakeholder relationships, and outcomes [Joh, 2014; Braga & Weisburd, 2015].

Researchers in sociology and public policy have also identified several technical and social challenges in the adoption and use of data-driven practices in policing. There are substantial IT infrastructure challenges that exacerbate IT management issues such as inaccurate data entry, lack of data quality, legacy systems, and siloed databases [Nunn, 2001]. Social issues around police big data use point out how data frequently perpetuates and confirms police biases since data is not objective in nature [Joh, 2014; Brayne, 2017]. For example, big data can reproduce inequalities by exacerbating surveillance for individuals already under suspicion. Similarly, researchers warn against relying on data alone to make decisions rather than supplementing independent police assessment. For example, researchers fear that data could prime officers to look for crime in certain communities, instead of neutrally assessing the risks [Brayne, 2017]. Big data in policing also magnifies concerns about the fourth amendment—the right to be secure against unreasonable searches—as it enables “perpetual indiscriminate data collection of entire populations” [Joh, 2014]. Similarly, data-driven policing critics raise concerns about the lack of understanding regarding long-term impacts of data-driven strategies, using data to justify over-policing disadvantaged communities, and community relationships with the police [Kochel, 2011; Gangadharan, 2012]. While new strategies create positive media attention, researchers raise concerns about the effectiveness of using big data products and practices from the for-profit industry in mission-driven organizations like policing [Sklansky, 2011; Kochel, 2011]. As mission-driven organizations, police agencies are often characterized by issues of hierarchical, rigid command structure of police organizations, lack of data dissemination through the ranks, and lack of accountability on how data collection feeds into their strategies and goals, which all adds to the data-related challenges police agencies face. Other challenges include police officers' resistance to change, preference for traditional policing, and difficulty implementing new programs [Capowich & Roehl, 1994; Sadd & Grinc, 1994]

Researchers recommend both technical and social efforts to alleviate these challenges for police in adopting data-driven policing. These recommendations include improving the IT infrastructures, aligning work practices with data-driven objectives, and providing committed management and expertise [Joh, 2014; Dabney, 2010; Manning, 2008]. This research contributes to this work by empirically identifying the social and technological challenges that hinder police accountability, as well as how stakeholders envision the design space around data-driven practices and tools.

2.3 Data Politics

Here, I review literature on how data politics manifests in public-services organizations as big data marks a computational turn in thinking within organizations. Data politics here refers to the inherent subjective and biased nature of data based on decisions around *what* is measured, *who* has the power to decide, and *how* the data is produced. Multiple researchers have raised concerns around the potential misuse of data in regards to how the politics of data manifest in organizations, such as: decontextualized nature of data, embedded epistemological biases, inaccurate or incomplete data framing narratives and space for action, and digital divides [boyd & Crawford, 2012; Berry, 2011; Crawford, 2013; Swenson, 2014]. This research investigates the problem space for mitigating these inherent biases and limitations of data use.

boyd and Crawford argue that 'big data' is an interplay of three elements: (1) technology that gathers, links, and analyses large data sets; (2) analysis for economic, social, technical and legal patterns; and (3) mythology that data can offer a higher form of intelligence and knowledge [boyd & Crawford, 2012]. Similarly, Morgan's seminal scholarship on the metaphors through which we understand organizations highlights quantitative data as one of the mythologies shaping organizational life, lending decision making a semblance of rationality [Morgan, 1997]. Although definitions of big data vary, researchers increasingly acknowledge that big data is less about the *size* of the data and more about gleaning knowledge from of the data [Burkholder, 1992; Ribes & Jackson, 2013]. Big data represents a social and cultural shift in how we create and use knowledge: “Big Data reframes key questions about the constitution of knowledge, the processes of research, how we should engage with information, and the nature and categorization of reality*”* [boyd & Crawford, 2012]. Because big data is a sociotechnical phenomenon, it entails all the biases that come from “human design” [Crawford, 2013]. There is, however, a dearth of empirical evidence of the nature of these biases and how they play out in practice.

While data can be misused due to its decontextualized nature, research also raises questions about the role of qualitative knowledge in decision making in data-driven organizations [Verma & Voida, 2016; Marshall et. al., 2016; Bopp et. al., 2017]. This concern is especially relevant for professions that require higher levels of *metis*, which refers to practical knowledge, gut instinct, and experience [Scott, 1988]. Metis is “knowing how and when to apply the rules of thumb in concrete situations” and is often grasped through experience instead of explicitly taught. As big data favors quantitative data, it is important to understand how this epistemological bias impacts organizations and their stakeholders [Verma & Voida, 2016; Marshall et. al., 2016].

Data is embedded with biases and values throughout data collection, cleaning, analysis, and presentation processes [Pine & Mazmanian, 2015; Verma & Voida, 2016]. While academics widely acknowledge data’s politicized nature, data-driven practices and technologies are marketed as enablers of objectivity, accuracy, and truth in order to provide accountability and legitimacy to decisions [Pine & Mazmanian, 2015; Suchman, 1993; Winner, 1999]. Values are also enacted in practice, through the use of the technology [LeDantec et al., 2009; Voida et. al., 2014]. And researchers have advocated for understanding values tensions in contexts where the same values may be shared by both technology design and end users but where the logics behind how those values are enacted are different [Voida et. al., 2014]. Subsequently, researchers call for empirical research on the nature of how these biases and data politics play out in implementation, especially how data can be used to frame narratives and shape the space for action [Crawford, 2013; boyd & Crawford, 2012; Verma & Voida, 2016].

Research about public services using data to support decision-making have also characterized issues of data perpetuating discriminatory actions such as algorithm-based technology recommending longer jail sentences for black offenders [Lum, 2017; Eckhouse, 2017]. Researchers have also warned about how marginalized communities can be at higher risks for harm and loss of opportunities through data inclusion by public services [Gangadharan, 2012]. Similar studies have found that data collection in public services is often marred with inaccurate accounts that fail to capture work flows and that logics behind organizational accountability can impact organizational effectiveness negatively [Pine & Mazmanian, 2014; Pine & Mazmanian, 2015].

Organizationally, big data brings issues of digital divides in organizations between those who have data access, skills, and literacy [Swenson, 2014; Ferguson et. al., 2014]. Researchers raise concerns about the biases of big data leading to new digital divides between data haves and have-nots and between individuals and organizations that do and do not have computational literacies [boyd & Crawford, 2012; Kitchin, 2014]. Manovich suggests that in this era of big data, there are three types of people: those who create data (both consciously and by leaving digital footprints), those who have the means to collect it, and those who have expertise to analyze it [Manovich, 2011].

This dissertation builds on this body of work to provide empirical accounts of how data politics impact police actions and accountability as well as investigate design directions for mitigating these inherent biases and limitations of data use.

2.4 Community Informatics

This research has indicated that simply adopting data-driven tools and practices does not inherently support accountability from stakeholder perspectives and it is important to explore how involving external and internal stakeholders in shaping data-driven policies and practices can address accountability concerns. In the follow section, I articulate key lessons learned from prior literature regarding the strengths and challenges when including such individuals into governmental policy and decision-making processes. First, I will discuss the strengths and limits of community lead governance as described in public policy literature. Second, I will look at how HCI, predominantly through community informatics, digital civics, and participatory design has informed design by collaborating with stakeholders.

Community engagement in governance provides formalized ways for citizens to participate in decision making around policies that impact them [Carol, 2007; Schuler, 1996], which is key for marginalized communities to have input in their local governance [Taylor, 2007]. Community participation in policy making allows access to a usually untapped expert source of knowledge from local citizens to stimulate collective action [McCabe et al., 2006; Ryan et al., 2006]. As community participation becomes a public policy agenda focus, several key elements have been noted for successful collaborations between the public and local government institutions, including a source of authority to legitimize public-government issues and practices, follow up evaluations, formal procedures and relationships to mediate community engagement, and demonstrations of practical engagement [Cavaye, 2004]. Success indicators for community initiatives include devolving decision-making powers, strategic reorganization, defining core values and commitments, ownership of initiatives, and establishing communication and rapport [Taylor, 2007; McCabe et al, 2006; Goodwin, 2005]. Barriers to successful collaborations include weak communication and feedback procedures, lack of transparent knowledge, lack of capacity and awareness within communities, resistance, under resourced initiatives, short-term planning, and rigid bureaucratic decision process [McCabe et al, 2006; Goodwin, 2005]. Subsequently, public policy research warns against the pitfalls of these barriers to the continuation of community initiatives. Often when community participation encounters obstacles, the initiatives for community engagement are theorized as failures. Eversole calls for ways to rethink community participation in governance beyond a single process but as “the juxtaposition of different ways of governing,” viewing community engagement through different lenses of varied capacities, skills, and roles in policy-making. This kind of renewed perspective on community engagement can create new possibilities for collaboration between government and communities, while recognizing the value and pragmatics of community contributions in theory and practice. Based on this public policy literature on community engagement, I transition to focus on how HCI has incorporated community participation in design.

Within the context of HCI, community informatics research points out the role of technology in supporting and creating opportunities for engaging in democratic processes [Muller & Kuhn, 1993; Carroll and Rosson, 2007]. Carroll and Rosson articulate the interplay of participatory design and community informatics with two propositions about why participatory design practices matter for community-impacted innovation [Carroll and Rosson, 2007]. The first is a moral one, where the people most impacted by a design should have a substantial say in how the system is designed, developed, and used. The second is a pragmatic proposition, where the people who will adopt and use the design should contribute their perspectives and preferences in order to improve the chances of a successful design outcome. Historically, participatory design stems from the demand for increased voice in decision-making processes with higher engagement from groups of members to represent shared interests and values [Simonsen & Robertson, 2012]. The underpinning goal is to provide users with better tools to support collective goals [Muller et. al., 1997]. Subsequently, researchers provide a framework of themes to outline high-level requirements for sustainable community informatics initiatives. The themes include identifying IT needs, organizing for IT change, learning new IT skills, and lastly, creating and sustaining intrinsic motivation for community participation. These requirements exist in tension with the general lack of resources (time, funding, skills, people) and the importance of direct participation in community initiatives. Researchers advocate to work *with* these barriers in order to not derail the projects.

HCI researchers have studied technology use in interactions between community members and governmental agencies [Erete, 2013; Lewis & Lewis, 2012; Erete et. al., 2014; Voida et. al., 2014]. Within in the context of community policing, we understand how online technology helps citizens unite and coordinate actions as well as regulate online and offline social norms [Lewis & Lewis, 2012]. In their study of a crime web forum, the researchers specifically call for technology fostering a space for collaboration beyond providing information. Erete speaks to the importance of including the marginalized communities that are disproportionately impacted by crime in policy-making around those issues [Erete, 2013; Erete et. al., 2017]. Her design recommendations advocate for elements of accountability, visibility, and participation to support community engagement. For instance, using online technology not associated with the police gave community members a space to discuss issues and solutions outside of the physical space to better organize as well as trace back action items. Technology that is not linked to the governmental agencies but fosters conversations around them helps alleviate the distrust community members feel. Erete points to the need of diversifying the avenues for participation due to personal, social, and financial barriers (*i.e.¸* second job, childcare, lack of transportation) and calls for more research on the role of technology in building trust and transparency between policing and marginalized communities [Erete, 2013].

Transparency about data production and how data becomes actionable is needed in order to establish accountability practices. Thus, transparency becomes a "pro-ethical condition" and not necessarily an ethical principle since it enables or impairs other ethical principles [Turilli & Floridi, 2009]. This work extends community informatics research by looking at where technology can supplement police-community relationships and community participation. Similarly, researchers have listed essential factors that support participation: access to relevant information, opportunity to take independent positions on issues, inclusion in the decision-making process, appropriate methods of participation, and flexible organizational and technological processes [Clement, 1993].

2.5 Accountability in Governance

Issues of power have remained central to government and politics, as philosophers throughout the ages have pondered and investigated the challenges of having too much or too little power [Madison, 1961]. How to keep it under control? How to prevent abuse of power? How to create checks and balances? Accountability is a manifestation of the continued concern about governmental powers and the need for oversight and surveillance into governmental actions [Schedler, 1999; Mulgan, 2000]. Political accountability is often defined and conceptualized through answerability—"the obligation of public officials to inform about their activities and to justify them"—and enforcement—"the capacity to impose negative sanctions on [those] who violate certain rules of conduct" [Schedler, 1999; Hales, 2008]. Subsequently, accountability is essential to democratic governments. As governmental organizations yield power over citizens, accountability mechanisms provide ways for citizens to assess, scrutinize and question the government's actions [Peters, 2007]. On an underlying level, these accountability mechanisms are assumed to support (a) identifying poor performance in order to (b) mobilize the public and ultimately, (c) create change in policies and practices [Peters, 2007]. In this section, I focus on understanding interpretations of accountability in governance and the limitations and challenges around these interpretations. Then, I specifically look at recent accountability efforts in the policing context in the US. Following that, I present HCI literature related to accountability in design theory and design practice and how this work fits in to calls for future work.

Since accountability's interpretations vary in governance, the way the term is used can have various implications for practices [Thomas, 2004]. Here, I outline three predominant and overlapping interpretations. The first interpretation of accountability focuses on transparency [Kaufman, 2005]. In order to support accountability, governmental actions must be transparent for "independent and external" reviews [Peters, 2007]. This includes answerability to specific concerns raised by stakeholders and general disclosure of unsolicited information about organization's behaviors and motives to stakeholders. While calls for transparency by activists and citizen groups have remained popular, research on accountability notes that there are gaps in *how* transparency supports accountability [Hales, 2008]. For instance, Hales notes "bad publicity" is not sufficient for government agents to change behavior, especially when "information users have no formal control over disclosers, and indeed may be significantly less powerful" [2008]. Unless accountability mechanisms include ways for stakeholders to "punish improper behavior," transparency mechanisms cannot sufficiently shed "light into the black box of politics" [Schedler, 1999]. For example, in February 2018, a local newspaper reported that MMPD failed to follow policy and review multiple police shootings over two years[3]. This raised concerns from the community about the futility of transparency if no one outside the organization has any power to create change. This research aims to understand how police and marginalized communities can mitigate the effects of these power disparities in making transparency based accountability more substantial.

The second interpretation of accountability relates to governmental responsibility, where government agents act accordingly and responsibly to the established law and ethics [Bovens, 1999]. In this case, accountability has more implications for an

---

[3] https://www.indystar.com/story/news/crime/2018/02/04/impd-violated-policy-failed-review-19-police-shootings/1082235001/

organizations' and workers' internal values and understanding of the law [Peters, 2007; Mulgan, 2000]. Researcher do distinguish between internal accountability within organizations (workers to management and superiors) and internalized accountability of individual workers (how and to whom they envision being accountable to *i.e.*, the public or governmental superiors). Previous scholars have raised questions about internal accountability in public services in terms of how non-expert supervisory members can hold expert professional to account [Day and Klein, 1987; Mulgan, 2000]. Similarly, research has also demonstrated that workers embrace internalized accountability with reluctance. One of the reasons the respondents reported was that they believed workers should be accountable to *someone*, but were not sure to whom exactly [Day and Klein, 1987]. Researchers raises questions and doubts about what kinds of accountability practices governmental agencies engage in and which groups of people are these efforts geared towards [Mulgan, 2000]. My research indicated similar tensions and disagreements amongst police officers about what kinds of accountability and to whom are appropriate as they navigate being data-driven.

Lastly, defining accountability can also refer to responsiveness of the governmental agents to the demands of their "political masters, clients, or the public at large" [Peters, 2007; Romzek & Dubnick, 1994]. Accountability here refers to the top-down, mission-driven nature of governmental agencies with civil servants being able to take directions from above, to attempt to serve the public [Mulgan, 2000]. Because accountability is about providing ways for different stakeholders to have voice in decision-making and policies, agents' ability and inclination to respond to demands from politicians and the wider public is key [Mulgan, 2000]. However, Mulgan points out that while responsiveness can be a dimension of holding agencies accountable, being more responsive does not equate to increased accountability. For instance, as some public services have adopted a more “client” or “customer” geared management approach to improving efficiency, organizations have become more responsive, or “customer friendly” but not necessarily more accountable.

Ultimately, while these three interpretations are not mutually exclusive, how organizations articulate accountability impacts how practices and mechanisms are adopted and shaped. This research specifically utilizes the first and second interpretation as lenses into understanding how community and police stakeholders, respectively, envision the role of data and technology in accountability. The first interpretation related to transparency and control to change practices is helpful in understanding community perspectives around police accountability. On the other hand, the interpretation related to responsibility is suitable in exploring how police envision accountability. It is not sufficient for police to provide more transparency through data if there are no mechanisms that allow communities to create change to practices and policies. Likewise, police’s data creation and use can continue to perpetuate narratives about crime and communities, which, besides being detrimental to marginalized communities, does not adequately address the calls for legitimizing the experiences of marginalized communities (*e.g.,* Black Lives Matter). This research aims to address these broader questions regarding accountability: how do we produce possibility of accountability when dealing with highly unequal social stakeholders in terms of power (police vs. marginalized communities)? How do we create spaces (socially and technologically) where accounts that are not formally included in police actions can be legitimized?

Next, I outline certain community and police-driven efforts around accountability in the US.

What about Police Accountability in the US?

Police accountability here is defined as a condition of being responsible and answerable to the public for *what* and *how* police perform on an agency and individual level, including how they do and do not equitably manage crime and disorder with respect to the law [Walker, 2001]. As the sociotechnical landscape shifts through the use and availability of social media and camera phones supplementing the rise of the Black Lives Matter movement, police organizations in the US are under intense scrutiny [Williams, 2015; Delgado & Stefancic, 2015]. This scrutiny is in part enabled by data, as critics of police use data to demonstrate the problems in policing [Lum, 2017; Eckhouse, 2017]. For example, Human Rights Data Analysis Group (HRDAG), a nonprofit organization, "applies rigorous science to the analysis of human rights violations" in the policing context[4]. HRDAG's US policing project assesses and improves upon police violence data accuracy. Their work is, in part, a response to the lack of systematic collection and aggregation of data about police violence in the country, which hinders police accountability [Lum, 2017; Eckhouse, 2017]. Similarly, activists have created projects like Mapping Police Violence, which is a web-based platform that demonstrates police violence through visualizations. Because the government does not have an aggregated data set about police killings, activists created a crowdsourced and comprehensive database about police killings, "searching social media, obituaries, criminal records databases, police reports and other sources"[5]. These kinds of projects

[4] https://hrdag.org/policing/
[5] https://mappingpoliceviolence.org

enable activists and marginalized communities to utilize technology and data science in order to create mechanisms and alternative spaces for holding police accountable.

On the governmental side, as a response to these criticisms, the US government has advocated for using data-driven strategies. For instance, Obama administration launched the Police Data Initiative (PDI) and Data-Driven Justice Initiative (DDJ) in 2015 and 2016, respectively, to use data to "increase transparency, build community trust, and strengthen accountability" [Davis et al., 2016]. As of late-2016, over 120 jurisdictions across the country, including my case site, had committed to both PDI and DDJ. The adoption rate of these initiatives reaches about 30% of the American population. However, the fate of these initiatives' development, implementation, and evaluation remains unclear under the current government administration.

Accountability in HCI

Accountability in Design Theory

Accountability as a concept and issue is pertinent to design. While traditionally in software production, accountability has focused on models of quality assurance, which is about measuring and evaluating the production processes and end products [Button and Sharrock 1998, Paulk et al 1993; Eriksen, 2002]. However, as Eriksen points out, accountability issues pertain very closely to social and organizational issues, and she calls for individuals and organizations to re-address accountability issues, keeping the "growing diversity and pervasiveness of technology" in mind:

> Considering the various ways in which accountability is referred to in the design community, and how different interpretations of the concept shade into and partially rely on each other, there seemed reason to pose them, at least as a starting point: Of what exactly is accountability an attribute or a feature? Is it always, inherently, 'good'? Who defines it? For whom? Under what conditions? [Eriksen, 2002]

As she juxtaposes accountability literature in HCI and CSCW with each other, she calls for designers to question accountability and for whom. Through the review, she emphasizes the importance of developing features to make processes and practices “visible and accountable.” Eriksen then introduces a “rough” tool referred to ‘figure of thought’ [Eriksen, 1998]. This mapping tool can be used to discuss complex concepts and relationships and map out work practices, interactions, formal and informal communications. Using a figure of thought with participants could help explore the intricacies of systems, as well as help with issues of “multi-perspectivity- the opening up for multiple voices in design” [Eriksen, 2002]. This research in ways responds to Eriksen’s calls for asking questions around accountability in terms of power and underlying goals in the policing context.

Similarly, Dombrowski’s work around social issues and design strategies highlights six goals of social justice: transformation, recognition, reciprocity, enablement, distribution, and accountability [Dombrowski et al., 2016]. The design strategy of accountability here refers to the ability to hold “responsible those who foster or unduly benefit from the oppression of others and identifying and assigning appropriate sanctions, penalties, or even punishments.” The researchers also acknowledge the conundrum of accountability: if accountability is a mechanism against the misuses of power, how can anyone truly hold the powerful accountable without having substantial social standing and power themselves? What does this mean for marginalized communities, where justice has been violated through systemic means? While there are no easy solutions to these larger questions, the authors outline “necessary commitments” for social justice oriented design practices, which include a commitment to conflict, reflectivity, and personal ethics and politics. These commitments aim to foster political responsibility within designers as social justice issues continue to emerge through our technology production and use. Finally, the authors also advocate for building cause-based alliances to cultivate political capital and action for holding individuals and/or organizations accountable.

Ultimately, the underlying basis for a lot of HCI and design research remains Suchman's seminal work, *Located Accountabilities in Technology Production* [2000]. Suchman advocates for 'located accountability' or recognizing the limited and partial nature of our "vision" as designers in order to take personal responsibility for technology production. This work counters the objectivist stance where it is not possible to locate responsibility for how technology is created by pointing out that designers must reflect on and be aware of their own position in the larger network of social relations. Accountability for whom as a question is then about embodying personal accountability in design practices.

As HCI researchers advocate for awareness and ethics around our role as researchers and designers, this research aims to contribute to understanding ways to cultivate accountability in and through the design process. This work feeds into design implications for designers and technologists to support ethical use of data-driven technologies and practices.

Accountability in Data Practice

As big data becomes more prevalent in society, it raises questions about the role of data in supporting transparency and accountability. In this section, I cover previous work on how decisions around data collection and use support accountability and for whom. Depending on the kinds of data collected, practices around data as well as available infrastructure and expertise, data can legitimize and rationalize certain organizational and individual actions. Referring back to the political nature of data, it is important to emphasize that data by itself or more data does not necessarily increase accountability [Verma & Dombrowski, 2018]. Researchers continue to indicate the need to understand *how* data supports accountability in different contexts, as well as explore alternative spaces and ways of using data to hold organizations accountable.

For example, Pine and Liboiron conducted a case study highlighting the politics embedded into measurements, techniques and interfaces [Pine & Liboiron, 2015]. They point out that while interfaces usually present data in ways that are devoid of any subjective "human element," politicized human-computer interactions begin before the data is entered into a system. They demonstrate the concept of politics and data through a case study measuring maternal morbidity. While maternal morbidity rates in America are severely high, they found that obtaining precise maternal death data was quite difficult due to how the cause of death is recorded; where pregnancy is the cause of death, the recorded data is often mischaracterized (*i.e.,* cardiac arrest). The measured categories do not support identifying a specific problem for a high-risk population. The authors directly relate the way this data is (or rather, is not) measured to larger political issues of women's care not being taken seriously enough by healthcare professionals and policy-makers. Their multiple examples indicate how data-driven efforts towards accountability do not inherently stand to improve social injustices. Rather, it matters what is measured, how it is measured, and ultimately, for whom.

Another example pertaining to research around data for accountability is Irani and Silberman's work around the ethics of crowd work, specifically looking at Amazon Mechanical Turk workers [2016]. In this design project, they deal with issues of worker invisibility and unfair compensation for crowd workers by designing an alternate system named Turkopticon. This system offered a space for workers to review and share their experiences with the employers, since Mechanical Turk only provides ways for employers to review crowd workers. By building a data collection and aggregation system to make crowd workers' experiences visible, the researchers introduced an accountability mechanism as well as a decision support tool for marginalized workers. HCI research around accountability and data practices demonstrates the gaps in limitations of data use and how data often becomes a tool to shape spaces for action, as well as noting the importance of evaluating data-driven practices and tools. With this study, I contribute a deeper understanding of the impact of being data-driven on police accountability through police and community perspectives.

Chapter 3: Methods

In this three phase research, I conducted case studies mainly using qualitative methods to better understand data-driven technologies and practices, related sociotechnical challenges, and possible design interventions for human-services.

In the first phase, I conducted an exploratory study of non-profit use of data through semi-structured interviews. This work enabled me to understand the nature of stakeholders' work, the data practices and tools they utilize, and how data use relates to the mission of the organization. I interviewed participants across the organizational hierarchy to explore the challenges they face while making decisions with data and how this type of decision-making stands to impact the clients they serve. This set of empirical work provides me with a better understanding and orientation of what technology and data use entails at a nonprofit, in terms of organizational constrains. This phase of research informed the deeper dive into how data use impacts human-services organizations and their stakeholders.

Next, during the second phase of this research, I built off this knowledge to understand another human-services context: law enforcement. While the nonprofit sector's use of data shed lights on how data-driven decision-making can be problematized, I wanted this research to be applicable to broader contexts. Law enforcement being a charged context is riddled with criticisms around excessive force, racial profiling, and unjust discretionary practices. Researchers and activists, alike, have called for more stringent investigations of how technology and data stand to shape police work and in what ways. Initially, the goal of this research was to study a Fusion Center (a Federal intelligence agency) and their data use, however, case site access was declined for privacy and security reasons. This led me to investigate further around how data use is handled in more local, law enforcement settings. In this study, I snowball sampled my participants, which included patrol officers, lieutenants, detectives, etc. from Midwestern Metropolitan Police Department. On top of the interviews, I also conducted participant observations in several types of police and community meetings and events. While my case site access was still limited in terms of what information was shared with me and which technologies I was allowed to know about and witness, these qualitative methods helped me to build a solid understanding and sensitivity to the challenges that the police face in their daily work, especially considering the intense scrutiny they operate under. These challenges when coupled with data use demonstrated how limited data use can be when dealing with human beings on-the-ground and the sociotechnical opportunities to improve police accountability and relationships with the police.

The final phase of this research continued the investigation of how data does and does not support police accountability from various stakeholders' perspectives. Because police accountability is a social, contentious topic with viewpoints across the spectrum, it became important to understand how concerns around policing and data use manifest across different communities, which are involved with and impacted by police actions. The interviews in this study focused on understanding how different stakeholders conceptualize "accountability" and how they see data and technology playing a role in this context. The interview sessions also included a brief design activity to help us explore the alternatives to current data use in ways that could possibly help mitigate the relationship issues between police and communities.

This dissertation is a product of these research activities. By utilizing these qualitative methods, I am able to identify the practical and on-the-ground limitations and challenges data use and how data-driven rhetoric is intertwined with larger principles of truth, accountability, and power. In the final chapters, I illustrate how these research insights can inform designer and practitioner work. In what follows, I elaborate on each research phase.

3.1 Phase 1: Mythologies of Being Data-Driven and Actionable

I conducted a case study of the use of business intelligence in one human services organization. Case studies are a powerful method for deriving in-depth insights in an organizational context [Kitchin, 2014]. Existing case studies of BI in the private sector have focused on characterizing challenges of and success factors for BI adoption [Haskins & Baron, 2011; Gartner, 2016]. Here, my focus is on the mythologies of BI use and the ways in which the design of BI systems supports or thwarts these mythologies.

Participants and Data Collection

I conducted semi-structured interviews (76 minutes on average) with 17 individuals (5 female) who have end-user licenses to use Domo and sometimes other BI tools for their work at Helping Hand, a large, local affiliate of a national nonprofit organization that assists low–income populations. 13 participants held positions in mid- and upper- level management across several departments of Helping Hand; 4 participants worked in the IT and BI departments and were responsible for the backend data warehousing and the front end data analytics.

I conducted semi-structured interviews with each participant using a protocol focused on the following areas of inquiry:

- The nature of the participants' work, their roles in the organization, and how the participants understood their work to fit into the mission of the organization;
- The different data sources they use in their work; how they collect, extract, analyze and explore that data; and the ways they make decisions with or without that data;
- The ways that the data they use relates to the mission of the organization—whether it supports or complicates the mission; and
- Their experiences of the constraints and benefits of business intelligence.
- The interviews were transcribed on a rolling basis to facilitate ongoing analysis.

I analyzed data iteratively and inductively using grounded theory [Strauss & Corbin, 1997]. The initial open coding foregrounded what a culture of data meant to our participants, resulting in thirty-three values-related coding categories. Through affinity diagramming and axial coding, I identified four core values: *data-driven*, *predictive and proactive*, *shared accountability*, and *inquisitiveness*. I returned to the data related to these core values, conducting another round of coding focused specifically on understanding the role of technology as it supports or thwarts these values, noting that these values also aligned with the mythologies ascribed to big data and BI tools. Through this analysis, I identified a series of disconnects between aggregate and drill down views of data that fundamentally shape and are shaped by understandings of what data is "actionable."

3.2 Phase 2: Challenges in Data-Driven Policing for Social Criticisms

I conducted a field study of the Midwest Metropolitan Police Department (MMPD) studying their efforts to be data-driven. Field studies have been proclaimed as substantially important in studying the police field because they allow for closer consideration of how citizen interactions in policing evolve [Yin 2017; Manning, 2005]. Here, I focus on law enforcement challenges in conducting data-driven policing.

Before I proceed, it is important to note the challenges in accessing participants and spaces around the context of law enforcement. Initially, I requested organizational permission to research with the local Fusion Center, which is meant to connect local, state, and federal law enforcement agencies through shared intelligence. The administration denied my access request. Then, I decided to pursue access with a local law enforcement agency, MMPD, but did not hear back from the administration about their decision. During this time, I also attended numerous police-community meetings in order to speak to police officers directly. In those interactions, police officers I spoke with were reluctant to sit down for an interview. They mostly advised me to reach out the police station. At that point around nine months had passed without getting any kind of access to police or the police organization. Then, a personal acquaintance offered to connect me with a community relations police officer they knew. After about two months, the community relations police officer responded back and agreed to meet with me as my first participant. From then, I was able to speak to multiple police officers through snowball sampling. The challenges in accessing a contentious space as a researcher shed light on how my personal identity has shaped my research. Being an outsider to the law enforcement context did not induce trust from the police officers I spoke with. Similarly, in the year it took to get access, I revised how I presented my research objectives to the police officers in order to help them see the value for them with this research (*i.e.,* improvements and support for their technological and data-driven practices). Subsequently, once I was in with a couple of the police participants, it became easier to convince other police officers to participate because they were aware that I have been speaking with their colleagues already. To demonstrate how aggressively I had to seek out opportunities to speak with police officers, I even conducted one of my interviews in a church parking lot spontaneously, after getting pulled over for speeding and convincing the officer to speak with me about my work. However, due to the limited nature of my site access, participants often spoke to me in a constrained, and sometimes, defensive manner. Throughout the data-collection, it was evident that I, as researcher, was an outsider to them, which undermined participant's candidness. Given this, I did not always explicitly discuss certain sensitive issues around police and their technology use (*e.g.*, specific details of their technology and data use, social criticisms towards police), depending on the participants' demeanor. It is also important to note that the participants I did recruit were mostly public-facing agents, who had a range of training around how to talk to the public around contentious issues.

Participants and Data Collection

I conducted 16 semi-structured interviews (98 minutes, on average) with MMPD members, including three civilians, three on-the-ground officers, three detectives, two lieutenants, two sergeants, and two captains across three of the six districts of MMPD. The civilian participants worked directly with data at the Information and Intelligence Center, while the sworn officers often held multiple responsibilities in terms of working with data and policing.

I conducted the interviews with each participant using a protocol focused on the nature of the participants' work, their roles in the organization, and how the participants understood their work to fit into the mission of the organization; the different technologies and data sources they use in their work; how they collect, extract, analyze and explore data; the ways they make decisions with or without that data; and the ways that the data they use relates to the mission of the organization—the rationality behind data-driven policing; and their experiences of the challenges and benefits of technology and data use. I continually transcribed interviews to facilitate concurrent analysis.

In addition to the interviews, we conducted 55 hours of observations, including community-police meetings, public safety meetings and training, and accompanying police during their patrols to understand their work's breadth and context. The observations were conducted from September 2016 to July 2017. These observations focused on who attended these meetings, their roles in the communities, what kinds of information and data the police shared, what kinds of questions and concerns were raised by the community members and police officers. For ride alongs with patrol officers, my observations focused on understanding police officers' workflow, the different ways technology in their cars facilitated their work, as well as the challenges and workaround police employ in conducting their calls.

I analyzed data iteratively and inductively [Corbin & Strauss, 2008]. The initial open coding foregrounded what a data-driven approach to policing meant to our participants, which lead to five overarching categories related to police's technology and data related challenges. Through iterative memoing and axial coding, I identified three sets of challenges MMPD navigate during data use as well as the role of data as it supports or thwarts police work.

3.3 Phase 3: Police Accountability and Data

In this final phase, I conducted a field study in the Midwest region with law enforcement, community and activist stakeholders. The focus of this study was to understand how different stakeholders conceptualize and experience police accountability and the challenges and opportunities of data use in mitigating issues around police accountability. While I was able to use my connections with the police from the second study to recruit police participants, I did face some minor challenges in accessing activists as well. This was primarily with abolitionist activists, who deeply question the legitimacy of the police. While these activists did speak with me, a couple of them were hesitant and one of them was somewhat aggressive around my research with the police. From their perspective, researching with the police in this research setting could be considered "complicit." However, in response to that critique, this work has ultimately been an attempt to study power up. As police officers who hold high levels of power in society, it is important to understand the challenges they experience in using technology in their work and the related sociotechnical implications. Studying the police and community experiences gives this work a more robust vantage point to discuss the problems around data-driven police accountability. Moreover, this final study gives voice to the power disparate groups in the conversations that shape data-driven policing accountability that are often not placed in conversation with the police. Subsequently, it is important to note that there were several activists, community members, and police officers in this study who expressed that they saw research like this as necessary and helpful in advancing the conversations and progress toward police accountability.

The restrictions around my recruiting in all these various groups was also apparent in what kinds of data collection activities I could engage in. While initially I had wanted to conduct focus groups with these various stakeholders, it became clear that without organizational approval, I did not have the leverage to schedule these participants to show up. Similarly, after initial research design, it was also apparent that several of the police and activist participants would not feel comfortable participating in a shared spaces, which is why one-on-one interviews provided me at the best chance of collecting data on their authentic experiences.

Participants and Data Collection

I conducted 16 semi-structured interviews, which lasted 92 minutes on average. Participants were recruited from the local law enforcement agencies, neighborhood crime watch organizations, and activist groups predominantly from the Midwest region.
I wanted to speak to groups of people that regularly engage with and are impacted by the local police, in order to understand the ways these perspectives shaped police accountability and the role of data.

The interviews were conducted throughout the city, at police headquarters, local coffee shops, or on campus. The interviews were conducted using a protocol focused on the nature of the participant's work, their role in their organizations and the community, their experiences and perceptions of police accountability, the different technologies and data sources they utilize in their work; the ways that data use relates to holding police accountable and the related challenges and benefits around such contentious and political issues. While the interviews all followed the same line of inquiry, I created three different versions of the interview protocols to address the underlying assumptions of police's legitimacy in these various stakeholders' eyes (see Table 1).

Table 1: Phase 3 Study Participants

Stakeholder Group	**Identify As**	**Participant Numbers**
Police* (5 participants)	MMPD	P1, P4, P9, P11, P12
Activist** (5 participants)	Reformer	P2, P15
	Abolitionist	P3, P8, P13
Community*** (6 participants)	Strong Police Support	P6, P7, P14
	Critical Support	P5, P10, P16

*Police: these participants were revisited from the second study, and included detectives, street officers, and lieutenants.
**Activist: these participants were active in at least one local activist organization around police brutality and race-related issues. While most of these participants are from and work in the Midwest region, one activist was based in Austin, TX. While all of these participants were highly critical of the police and their practices, three of the five participants described themselves as abolitionists, who do not believe that the police are a legitimate and needed part of society because of the history and current practices of systemic and embedded racism. These participants were recruited from local activist meetings as well as through snowball sampling.
**Community: these participants were recruited local Crime Watch chapters and police-community meetings. I wanted to get the perspectives of community members, who choose to engage with the local police to both understand their motivation to engage with the police as well as how they perceive issues around police accountability. Three of the six participants were staunch police supporters, whereas the rest were a bit more critical of the police, but not completely opposed to the police's existence.

Design Brainstorming Activity

Towards the end of the interview session, the participants were also given two scenarios (see Appendix) depicting the limitations of data use in police accountability work. These scenarios were developed in consultation with two police officers and two activists. The first scenario focused on demonstrating how data fails to depict police activity and impact in a meaningful way from the police perspective. The second scenario illustrated how decisions around collecting and using data are inaccessible to the community, raising questions around how data use perpetuates existing inequalities for marginalized communities. These scenarios were used to spark conversations and brainstorming around possible, future uses and practices around data to support principles like accountability. I continually transcribed interviews to facilitate concurrent analysis.

Data Analysis

I analyzed data iteratively and inductively [Corbin & Strauss, 2008]. The initial open coding foregrounded what police accountability means to different participants, which led six overarching categories related to how data stands to be a function of power. Through iterative memoing and axial coding, I identified three categories of ways data use perpetuates power disparities for these groups of participants. These categories were then related to how my participants saw data supporting more accountable futures around police. Because the sample for this research was comprised of a heterogeneous group of stakeholders with varying backgrounds, jobs, and agendas, analysis of this research focused on how police accountability was specifically experienced in various ways by different groups of people in society. Consistent themes around police responsibility, responsiveness, and justice rose through this analysis, but with a highly disparate range of issues for different participants. I analyzed the different perspectives on these themes by memoing how stakeholders experienced issues around power and social disparities. Ultimately, the analysis was conducted with a focus on how stakeholders with highly unequal power disparities and social standing experience and navigate data-driven policing, as well as the possible design spaces to mitigate those disparities. Because of the research goal of understanding and mitigating existing power disparities, this analysis focused on giving voice to stakeholders without the same level of power in data-driven policing.

Chapter 4: Mythologies of Data-Driven Technologies and Practices

4.1 Research Context

Helping Hands[6] is one of a relatively small number of nonprofit organizations using data-driven tools such as business intelligence (BI) technology, and the particular data needs and pressures in the nonprofit context highlight challenges in how data-driven work is handled. Helping Hand is a large, local affiliate of a national human services organization that assists low–income populations through a range of programs and services:

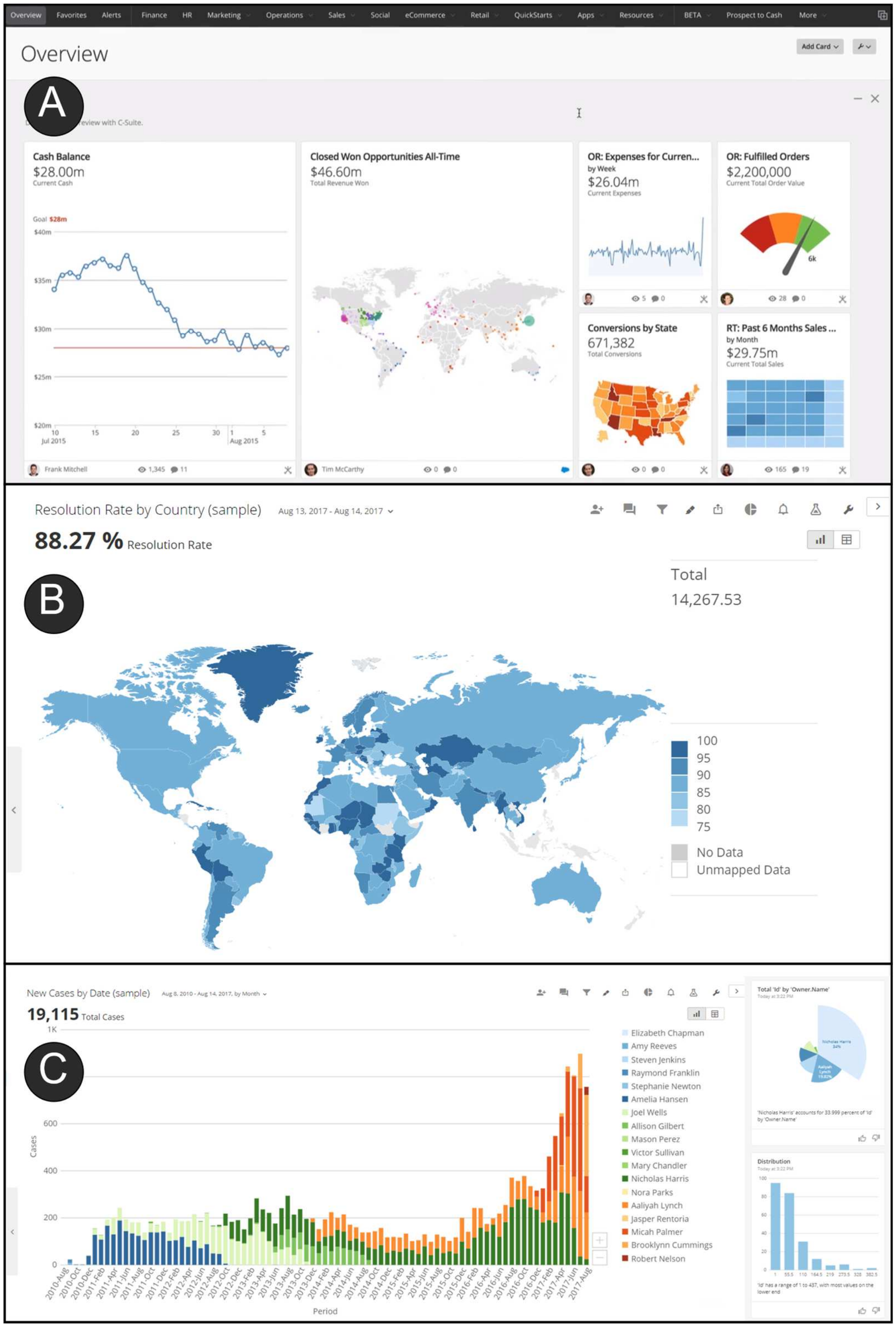
- **Business Services:** Helping Hand's business services department operates small businesses with employees who are often clients of the organization. The information management needs of this division include employee scheduling, production, inventory, and revenue.
- **Mission Services:** Although all departments operate within the same, overarching mission, the mission services department offers a variety of programs to support the resiliency of their clients. This department relies on information management to understand the *impact* that the organization has on clients. Significant information challenges center around questions about how to assess the impact of its programs and what information should be collected to do so.
- **Education Services**. The education services department at Helping Hand manages charter schools for low–income, at–risk youth. Their information needs include demographic information about their student population, class scheduling,

[6] The name of the organization and all of its internal departments have been anonymized.

records of student attendance and achievement, job placement, and salary of placement.

In early 2013, Helping Hand was awarded a small grant to fund the purchase of 50 licenses for the data-driven system Domo, as well as salary support for a business intelligence staff position. The primary motivation for the adoption of Domo was to promote a “culture of data” within the organization, to support their actions “using sound evidence” [P12]. Since the goal behind data-driven tool adoption was to aggregate data across several departments and systems to enable a centralized view, the BI Manager held numerous individual meetings and focus groups with various organizational stakeholders to identify and prioritize key metrics, start to wrangle data from across a breadth of sources into their data warehouse, and coordinate end-user training. Based on usage log data at the time of the interviews in the fall of 2014, Helping Hand estimated that of the 50 licenses that were purchased, Domo had 15 daily or weekly users; 15 monthly users (accessed primarily for monthly reporting activities), 5 users who had not logged in since their initial training; and 15 users who were still waiting for their data to be added to or configured in the backend data warehouse.

Figure 1: Domo Dashboards



Note. A: Dashboards in Domo are customizable to provide snapshots of aggregated views of data. B: Aggregate views in Domo provide high-level visualizations of data (e.g., Resolution rate by country). C: From the aggregate view, users can drill down into quantitative data with more granularity.

Each Domo user accesses data via a dashboard, initially setup by the manager of BI (e.g., Figure 1a). The dashboard is tiled with "cards" that represent a query performed on the data at different levels. There are multiple levels of granularity that can be offered as filters and thresholds are customizable. Clicking on a card reveals an aggregate–level view of the data resulting from a given filter or query, offering high-level, trends and patterns analysis (e.g., Figure 1b). Clicking further on the aggregate-level view accesses the drill down, which provides more granularity to the quantitative data (e.g., Figure 1c).

Domo is currently used at the highest levels of Helping Hand management and across the leadership of all departments. Other BI software is used—with varying degrees of success and varying degrees of redundancy—across different subsets of departments. Education services, for example, also uses *Tableau*. As suggested by previous research, participants in all departments used their BI system(s) as one small part of a broader ecology of information management tools—using various Excel spreadsheets, Outlook Address books, and paper-based systems to accommodate the needs and individual styles of their knowledge work [Kidd, 1994; Voida et. al. 2011].

4.2 Results

In the following sections, I introduce four core mythologies: *data-driven*, *predictive and proactive*, *shared accountability*, and *inquisitiveness*. Participants most frequently experienced these mythologies in terms of organizational values, both instrumental and terminal [Hannula & Pirttimaki, 2003]. These mythologies align with the common marketing of big data and business intelligence. Yet, through the discussion of each mythology, I highlight the ways in which the enactment of each is problematized by recurring disconnects between aggregate views of data and its drill-down in business intelligence systems. These disconnects relate to participants' understandings of what it means for data to be actionable and valid for data-driven decision making.

Note that the participants frequently used the phrase 'drill down' both in a literal sense—to use the drill down feature in the BI tool to get finer granularity quantitative data—and also, more commonly, in a metaphorical sense—to get more information that does not actually exist in their BI system. I use the term 'drill down' in the same multi-faceted fashion.

Data-Driven

The upper- and middle- level management participants at Helping Hand all speak positively and optimistically about the organization's ability to use data moving forward to improve program strategies, personnel evaluations, and workflow to serve their overarching mission better than before. Yet, participants have differing and sometimes conflicting perspectives about what kinds of data should be considered legitimate for substantiating the organization's impact and/or actionable for decision making.

Most participants conveyed a significant inclination towards using quantitative data to "prove" the effectiveness of their individual performance or the impact of the organization's work. For these participants, quantitative data is seen as the only acceptable indicator or "picture" of performance for many stakeholders:

> It's really a prove-it-to-me type of mentality and I think its data that's going to help us do that. [P13]

> So that's what we are going to try to use to the data to really drive us, and you can't quantify everything, that's just the reality, we are aware of that… but it does paint a pretty nice picture. [P3]

From the participants' perspectives, quantitative data "proves" impact whereas qualitative data helps people "connect emotionally" to the mission of the organization:

"We want them to… connect emotionally to what we are doing. So… we tell specific stories" [P17].

Most participants do not explicitly point to qualitative data as a legitimate basis for data-driven decision making. For these individuals, qualitative data-driven decision making isn't authoritatively substantiated, it is "just based on… anecdote" [P12]. Only one participant articulated a view of qualitative data as being a "somewhat" legitimate form of empirical data for serving as the basis of data-driven decision making. He describes qualitative data as "observational" and "unstructured":

> There is data input to every decision but some of it like I say is not in digital form, it's just observation, observational data… So that's somewhat data driven but [it's] unstructured data. And so that generates another conversation or decision point. [P6]

Here, though, the strongest hedge in the participant's language isn't related to the qualitative nature of the data but the fact that the data is not in digital form. He describes the use of observational, unstructured data as actionable because it enables him to take action, for example to have a follow-up conversation. Other participants questioned whether data had been appropriately vetted and whether they were a reasonable basis for communication and decision making if they were not included in Domo.

Although most participants did not explicitly identify qualitative data as being a legitimate basis for being data-driven, nearly all participants recounted experiences of data-driven decision making that centered around the use qualitative data. P11, for example, reflected on an instance in which he wanted to troubleshoot production issues and expressed frustration that the data available to him in the BI system lacked the qualitative, "human element" that he wanted:

> I can see if, you know, you're missing [production] because you don't have enough people producing or the people you do have producing are

> producing at half. And then when I am, you know, coaching… there is also that human element: "Well there was a death in that family, you know, I lost two people, I haven't been able to replace them yet. I'm working on that." That, you know, usually good reasons behind it and they're addressing it and they get right back up but, you know, if I could drill a little bit deeper. [P11]

This participant explains how quantitative data is used for keeping track of their production, but he wishes he could drill down "deeper," beyond the quantitative data in the system to qualitative data that could explain the context behind the numbers. In order to "drill down" to the depth that he needs, this participant has to speak to the site leaders to incorporate the "human element" into his understanding. The "drill down" data that he seeks is not actually captured in the BI tool—nor could it easily be given the quantitative emphasis of existing BI tools. Here, the BI system seems to exacerbate the uncertainty that the participants experience in considering whether qualitative data are a legitimate basis for being data-driven.

Among the participants at Helping Hand, Domo is held up as an embodiment of the promise of the data-driven organization, particularly as it represents the aggregation of their activity: "At the core of our approach, one of our central tenets is measuring outcomes with data and with this system, Domo, that aggregates everything we do…" [P5]. Yet, the aggregation of data in Domo supports only part of being actionable in their work. This same participant continues to emphasize the complementary need to "drill down" to the context surrounding the individual clients who are being served. And, he emphasizes that the aggregate views of quantitative data are most valuable when they are used in service of the "drill down" views of data that, ideally, enable them to understand why an individual has been successful or not:

> Our ability to measure outcomes dramatically affects how we can serve an individual. So knowing across the board where we are successful

> generally speaking and being able to drill down and look at on an individual basis how that success came about and we can do that. And that supports our mission in everything from the heart to the wallet… [P5]

For these participants, the relationship between being data driven and being actionable manifests through the conflicted interplay between quantitative and qualitative data. The language that the participants are almost uniformly using to characterize the relationship is a metaphor borrowed from the business intelligence tool they use—a relationship between the aggregate and the drill down. Yet, intriguingly, while the participants speak of the drill down as ideally providing qualitative, actionable evidence of the context surrounding the quantitative data, Domo (as with nearly all analytics tools) only provides quantitative "drill down" data, the individual-level quantitative data that is the basis for the aggregate-level quantitative data.

While the performance and legitimacy of the organization is supported by aggregate views of quantitative data, being actionable is supported by individual-level, qualitative "drill down" data that is important for responding to the unique circumstances of individuals. The participants use both qualitative and quantitative data but are unsure whether qualitative data is considered legitimate since this qualitative data doesn't actually exist in the business intelligence system. And it is frequently not even found in digital form. As such, its legitimacy is—at best—contested; at worst, the validity of this data as an actionable basis for decision making is threatened.

Predictive and Proactive

As participants work towards achieving a "culture of data" [P1] within their organization, they envision that a predictive use of data will also enable them to be more proactive. As they ramp up their business intelligence efforts, one participant characterizes the trajectory towards being proactive as the "real value" of these systems:

> I think we are on the cusp… of shifting from a reactive look at data to a predictive data…. The real value comes… when we can actually start to predict things that are going to happen and then intervene before they do. [P8]

According to this participant, the "real value" of business intelligence doesn't come just from the ability to aggregate historical data, but from being able to predict what is going to happen so he can act on it. This predictive capability, he continues, comes from being able to compare the drill-down "characteristics" of individuals with aggregate views of data, asking questions like: "What are the common characteristics of people who have graduated? Who have dropped out?" [P8]. Yet, to act on these aggregate views of data on behalf of individuals means walking a fine line between capitalizing on the predictive capabilities of the BI system and respecting the lived experiences of their clients:

> If we say that if you are an African American male, that's 23 or under, who has two kids, you know, who comes to us with fewer than 10 credits, you are highly unlikely to graduate. Right? It doesn't mean that the next African American male that shares these characteristics is not going to graduate but what we can do is start to surround him with additional support early to raise his chances right? So it's… it is profiling... but it's what we hope is profiling in a really, really positive manner. [P8]

The participant recognizes the disconnect between the quantitative, aggregated data and the individuals with real relationships and struggles that stand to be singled out but also surrounded with additional support as a result of predictive analytics. Despite the recognition of this uncomfortable disconnect, and without clear answers about the right path forward or the right language to use to describe the proactive work that is likely to happen at the drill-down, individual level, participants are still keen about the proactive use of predictive data to guide their actions as they serve their clients.

Shared Accountability

For the informants at Helping Hand, a "culture of data" should foster shared accountability among individuals across the organization as well as with external funders and community members. Since the mission of all their departments is to assist low–income populations, one informant points out how important it is to keep all the programs accountable to the shared organizational mission:

> My job is to use all of these programs plus all the resources that exist in [Helping Hand]… using sound evidence based programming…. It is all these contributing. We have to work together or it doesn't work. [P12]

The value of shared accountability is manifested in many of the informants' work practices, but most significantly in data reviews:

> Data reviews we started because, it's actually fundamental to, I think, the model. We want everybody to be accountable for their own data and to understand their own data. [P12]

Most informants view data reviews as an opportunity to address their performance and any issues associated with it in a transparent manner with other members of the organization, providing some additional context to the quantitative data.

While the shared visibility of data and the open discussion in data reviews may enable valuable forms of professional facework, it may also foster competitiveness within some subcultures internal to the organization:

> They are hyper competitive… and they're like one of the most data crazy groups that you ever see… so like data hungry. [They] will go through and say, "Well, I had this percentage of my students earn credit this past term year and you only had 10% lower than I did…" And I'll sit in meetings and they'll just totally call each other out… It's crazy! [P4]

From a management perspective, access to aggregate views of data also enables the leadership to identify outliers in the productivity of their workforce and coordinate mentors and other resources to help address whatever productivity gap might exist:

> If we see an issue at one school and success at another, we can say to them, the school leader, that's struggling, "Hey, you need to go talk to the school leader who is knocking it out of the park with that; and let's do some coaching there. [P8]

Although the focus here is on the employee rather than the client, there is the same emphasis on understanding and supporting the individual who underlies the data. Here, aggregate views of data are helpful in comparing employee productivity, but the drill down views of quantitative data are more deeply understood and acted on in conjunction with extensive qualitative data provided through the mentoring process, outside the BI tool.

But as with the multiple tenors of data use that emerge from shared accountability in data review meetings, there are also multiple tenors of data use resulting from the shared accountability of data with management. The middle- and upper- level management participants raise questions about how data should relate to employees' incentives and evaluation. Here, "performance" is used to reflect the more qualitative or subjective perceptions about employees' work whereas "outcomes" are reflections of work that have been metricized for the business intelligence system:

> I read a quote the other day…. It basically said something to the effect of… you cannot connect pay with performance because performance is circumstantial. But I think you can connect pay and data and incentives to outcomes… right? So performance and outcomes, I think, are different... Our perceived performance of something… our perception of somebody's performance could be totally different… but the outcomes could still be great. Or my perception is that the performance is great but the outcome is horrible…. And so that's what we'd like to do, is really make sure however you decide to achieve your goal… [P3]

This participant is still wrestling with the sometimes-conflicting forms of data that he receives about employees' work and acknowledges that observational data might not align with the quantitative data in the system. But he is still optimistic about finding some evidence-based means to evaluate employees against their goals.

Here, aggregate views of data enables the management to identify employees whose productivity levels are outliers. The disconnect in being actionable based on aggregate views of data stems from the lack of additional, contextual information in the BI system, which the participants then have to seek elsewhere. It is the conjunction of the quantitative data within and qualitative data outside the BI system that enables participants to be actionable. More generally, the mythology or desired value of shared accountability raises questions about how different individuals with access to data treat the human who underlies the data.

Inquisitiveness

As more data is integrated into the data warehouse and as more users have access through Domo to the data that they ideally want to use to make decisions and act on them, the participants hope that the system will enable them to be more inquisitive about the data. A few participants reported already having conducted hypothesis-driven mini experiments by studying aggregated, longitudinal data for certain trends. In one instance, the participant created a card in Domo to "prove" the effect of missing quota on production levels:

> I have a rolling twelve month card that runs production along with the sales… and the reason why we did that was because our production… was super low and we weren't making quota. And we were trying to prove to the regional managers, well if you make quota, the next week immediately your sales are up. [P1]

The readiness-at-hand of the data, in this case, empowered this user to ask questions of the data that they were curious about. Another participant discusses a similar hypothesis-driven study of data to answer his question about whether more communication about the mission of the organization makes their customers want to donate by rounding up their payment at the stores:

> One thing I'm curious about is that the stores that have that increased communication, are they… are people rounding up any more frequently at those stores because they're theoretically learning more about the mission than they would at stores where we don't have those communication efforts? The reason why I'm interested in that is because there could be a couple of different hypothesis on that… This data can help us to… prove that one way or another. [P13]

The ability to act on a value of inquisitiveness, however, relies on a certain level of technical and information literacy. Approximately 75% of all cards seen by all but one users are created by the IT or BI staff, who identify data sources, select a visualization widget, and configure the scope of the visualization. The two example mini-studies described earlier both rely on data presented in cards that had already been pre-configured in ways that were suitable for the questions they wanted to ask. Individuals who create their own cards have done so only after requesting and receiving multiple hours of one-on-one training from the IT or BI staff.

Creating a new card, however, requires some degree of scripting skills. At the time of the study, only one user had created his own cards by modifying the scripting from existing cards; the BI staff is unsure whether his cards have been configured correctly. If inquisitiveness persists as being an organizational value, it is one that likely will privilege users who learn new skills to support the dynamic creation of new cards to answer new kinds of questions.

Users who do not yet have these skills or who prefer to explore their data in other ways—the majority of our participants—either collect data in, sometimes redundantly, or export data to Excel spreadsheets. They feel it enables a richer and more accessible set of features for sense making than the drill down that is the sole analytic feature possible given a pre-defined set of cards:

> I can’t look at this [card] formatting. For me, I find it too hard. I think this is my problem, not the system’s problem…. If I am not in control of the columns… it’s too hard to look at, so I reformat everything…. [P17]

Most of the participants noted that their BI tools, including Domo, do not provide sufficient or sufficiently accessible control and flexibility for exploring and understanding their data. Even with a general understanding of the affordances of business intelligence tools, there is still a perceived disconnect between the resources and expertise required to make use of the pre-defined aggregate views of data and the dynamically-explorable data, ideally something beyond the drill down. In order to be inquisitive and ask questions of their data beyond the visualization widgets currently set up in their dashboards, the BI tools assume both scripting and data literacy skills beyond the current expertise of these users. Fostering inquisitiveness and supporting sensemaking through different drill downs is beyond the scope of *accessible* features for the majority of participants at Helping Hand.

4.3 Discussion

Mythologies of Data-Driven Systems and The Space for Action

Morgan’s work on organizational metaphors refers to quantitative data as one of the mythologies shaping organizational life by providing a semblance of rationality to decision making [Morgan et. al., 1997]. He claims that quantitative data in formal organizations plays the same role as magic in primitive societies, enabling clear-cut decisions to be made in situations that might otherwise be open-ended. Even though these techniques don’t reduce risks, the mythology of rationality as supported by quantitative analysis provides credibility to organizational actions. Similarly, the mythology of big data is believed to provide higher levels of intelligence with an aura of objectivity, truth and accuracy [boyd & Crawford, 2012]. These mythologies surrounding data compel us to question the values and biases that are embedded in organizational data and to critically examine the data that becomes legitimized through organizational action—what data is collected, what data is digitized, what data is aggregated and visualized in business intelligence systems [Crawford, 2013; Manovich, 2011]. It also compels us to question what kinds of action it may support or thwart.

If users consider the data in business intelligence systems to be the only valid representations of organizational ground truth for publically admissible data-driven decision making, as data reflecting participants' uncertainty about qualitative and unstructured data suggests, these biases stand to propagate through their actions. Just as the interplay of inclusion and exclusion of data in measurements can create a space for possible action [Crawford, 2013], the space for action within an organization can be constrained by the data and visualizations contained in the business intelligence system. Especially for a human services organization serving at-risk individuals, it is important to question what data is included and excluded from measurement to understand how the values embodied by data shape rational action and organizational culture.

4.4 Conclusion

For the participants at Helping Hand, the mythologies of business intelligence are experienced as powerful commitments to a set of organizational values. But as they attempt to enact these values through the use of BI tools, the full complement of data that they need to translate data into action are not supported by their information systems. And when data are not in the systems, there is clear uncertainty about whether data "counts" as a legitimate basis for data-driven decision making. Just as workflow systems were found to overconstrain work practices in organizations, this contemporary class of system also overconstrains work practices *and* ways of thinking about the work [Winograd, 1987]. The mythologies of business intelligence scope data in and out of the system, scope understandings about legitimacy, and scope the actions that are made based on data.

In this research, I have made the following contributions:

- Identified four core mythologies that characterize an organizational culture of data: data-driven, predictive and proactive, shared accountability and inquisitiveness;
- Identified breakdowns in data-driven decision making that stem from disconnects between the aggregate and drill down views of data in data-driven systems;
- Provided empirical evidence of the epistemological of data-driven systems propagating into confusion about what data is and should be considered legitimate for data-driven decision making; and
- Offered the first case study of the use of business intelligence and data analytics in a nonprofit organization, highlighting tensions in BI use that arise from the human-services context.

This empirical evidence suggests that the enactment of mythologies surrounding a data–driven culture require more comprehensive support for diverse types and combinations of data than are currently supported by this organization's ecosystem of information management tools. For the participants at Helping Hand, when they "drill down," they want to understand the "human element" represented by the data and they rely on that human element to help them know how to translate data into action. Given the recognition that there is a human being who underlies data, the question of how to act becomes a fundamentally moral one. And the design challenge we face is to re-envision the ecology of information management systems in ways that enable organizations to legitimize data that is most meaningful for being actionable, where what it means to be actionable may very well hinge on the moral treatment of the individuals who underlie data.

Chapter 5: Challenges and Limitations in Data-Driven Policing

5.1 Research Context

MMPD is the law enforcement agency of one of the most populous cities in the Midwest. Their mission focuses on reducing crime and the fear of crime while enhancing public safety. MMPD has six service districts, with about 1600 police officers in the workforce.

While MMPD has employed crime data analysts for over a decade, in early 2016, under a new administration, MMPD doubled down on efforts to be data-driven. This included revamping the Information and Intelligence Center, which employs different kinds of analysts (crime analysis, GIS, internet crimes, social media, real-time camera surveillance, and Uniform Crime Reporting (UCR) for the FBI). The Center is where requests are made for information for supporting ongoing investigations, writing funding grants, and developing new initiatives. Community members also send requests for city's crime data.

Through interviews and observations, my participants considered MMPD to be in the "infancy" stage of adopting data-driven policing. These participants echoed that their efforts towards data-driven policing were hindered by the lack of budget for technological and human resources. MMPD, at the time of our study, employed only three civilian data analysts, while most of the other analysts in the Information and Intelligence Center were sworn officers, who were either interested in technology or on "light duty" because of an injury, or held multiple positions in the organization. Though MMPD does have a small in-house IT department, an external, third-party agency controls and maintains their IT infrastructure. Several of my participants expressed dissatisfaction about how the external agency handles the police's technology needs because of incompatible IT best practices, lack of control over technology decisions, and organizational politics.

MMPD has four key data-driven practices relevant to their work, which I will discuss in turn. First, Compare Statistics, more commonly known as CompStat[7], is an adopted framework developed by New York Police Department for strategic problem solving and increased accountability [Weisburd et. al., 2003]. At MMPD, district commanders present detailed crime reports about their districts (provided by the analysts) to the chief of police once a month, where the administration makes decisions about resource allocations and ways of addressing crime. Second, Social Disorder Index (SDI) is part of the data used during CompStat meetings that guides decision-making. SDI divides the city into small grids (250 ft by 250 ft) and maps crime and other emergency incidents (*i.e.,* fire department, emergency medical services, high-risk SWAT areas). These incidents are then each given scores based on the gravity of the incident (*e.g.,* crimes against a person or property, mental illness, drug overdose). SDI is meant to help the police administration identify and address underlying patterns in high crime areas. Third, "heat list" analysis includes a running, monthly list of repeat victims and offenders in the districts. Lastly, community programs are implemented based on data about the area's needs. For instance, by looking at where the most juvenile crime occurs, MMPD was able to choose a location to organize a youth basketball league to provide better community interactions with the police.

[7] https://compstat.nypdonline.org/

MMPD's data collection includes, but is not limited to, computer-aided dispatch system, crime reporting system, license plate readers, closed circuit (CCTV) cameras (including collaborations with private businesses), social media, geographic information systems, and in-house forums for information dissemination. MMPD also collaborates with the local Fusion Center, which is a federally operated data exchange center to aid in police investigations.

5.2 Results

In the following sections, I identify three key challenges centered on MMPD's data-driven efforts: *data-driven frictions, precarious and inactionable insights, and police metis concerns.* These challenges shed light on how MMPD navigates new work possibilities afforded through data collection and analysis and balancing these new data forms with the human-element in police work. Through the discussion of these conversations, I highlight how data-driven challenges problematize police work.

Data-Driven Frictions

Within MMPD, there are two approaches regarding how technology and data can mitigate pressing social concerns. The first data-driven approach supports officers by creating and using data to help legitimize decisions and position practices as arguably objective and unbiased. The second data-driven approach modifies existing police practice towards the tenets of community policing, by not focusing on solely enforcing the law, but by using data to provide a proactive policing orientation. Data is used to identify underlying issues fostering crime in order to attempt to address those issues through community collaborations. For most of my participants, the first is the current status quo, whereas, the second approach is an idealized vision some police want to progress towards. In the following section, I discuss the conversations policing agents have around the role of data and how these technological visions impact police practices. Here, police officers face challenges with new forms of data that offer new police goals, however, because of their limited resources, they experience tensions in choosing which agendas to actively work towards.

Data for Legitimization Measures

As MMPD works to understand and address criticisms about perpetuating social inequalities, police officers are thinking through their own work and role. In my case study, police agents predominantly use data to legitimize existing practices and decisions. For instance, a patrol officer in our study echoed a widespread sentiment about data collection efforts as a mechanism to “CYA”3 (i.e., cover your ass) in case there were any complaints filed against the officers and their work. For instance, during my observations, a police officer made timestamped comments within their dispatch system on the laptop after a store manager called about theft. The officer explained her documentation as a way for her supervisor to retrieve the officer’s side of the events in case a complaint by the store manager was made against MMPD. Similarly, another patrol officer would mark their location while driving through areas where community members had requested more patrol in order to provide data about the numbers of police officers in that area.

Police also use data to garner public support for their practices and actions because of data’s perceived objectivity.

> …with [data] we get more buy-in, we get more support…. I think that's an effective result of data because again the community they get objective statistics from us now. It's more transparent for them [community members], that's one of the goals. [P6]

> … one thing that [the police department] hope[s] to do with data is to have it more readily available for the public when incidents happen…to readily share accurate data with the public is crucial sometimes in keeping that transparency and accountability with the department. [P4]

As police begin using data as a mechanism for transparency and accountability with the public, this raises questions about how and who collects, analyses, and presents data. These questions about what the data does and does not depict have to be addressed in order to better support efforts of improving police and community relations. Likewise, more data does not necessarily equate to more transparent and accountable police forces—for instance, while initial reports suggested bodycams influenced police to be less aggressive with the police and cause a decline in reports against the police [Ariel et. al. 2016; Stanley, 2013], research also showed the ways in which police shape recordings (i.e., shutting off cameras, placing evidence) [Ariel et. al. 2015].

Data for Proactive, Community-Oriented Measures

Next, I demonstrate how the proactive policing approach is part of the data-driven conversation at MMPD. Most of the participants from the Information and Intelligence Center envisioned using data to proactively address social issues and to change on-the-ground and administrative practices. One of the data-driven practices implemented by MMPD is the SDI, as introduced in the Research Context. Rather than just enforcing the law, the SDI aims to help police understand the prevalence of criminal and social issues across the city. In particular, the tool is meant to help the police administration identify and address underlying issues of high crime areas.

> You have to have all three things to have the crime occur. […] if we take any part of that convergence triad away, whether it is education for the victim, identifying the suspects early [and] making sure they are not in a position to re-offend, or changing the environment then we do not have the crime… their goal is to move forward to that predictive end. Right now, we're really providing what has happened in the past… [the data]

> starts building a better bigger picture of how is that convergence setting created. [P12]

For several MMPD officers, being data-driven enables them to understand what comorbid social conditions may underlie the city's criminal patterns. Besides basing resource allocation decisions off of data, community outreach programs created based on community needs data. Community outreach at MMPD includes programs for at-risk youth to provide positive role-models, fundraising efforts for less advantaged communities, conducting community meetings with newer immigrant communities to develop trust, and helping community members develop evidence-based neighborhood safety strategies. Similarly, the repeat victim and offender heat lists enable officers to focus patrols. These data-driven practices, help officers attend to a larger range of issues to shift police's focus towards preventing crime rather than just enforcing the law.

> …you're not going to arrest your way out of [crime], and we think, in our case, certain areas of crime are related to poverty and mental health. [A police employee] has access to all social services so when we go in and take a bunch of guys to jail, they go in and say [to the remaining family], 'we know your primary breadwinner, although [through] illegal [activities], is gone, can we help you with food and rent and social services'… To connect social services with those people, who would retreat to crime otherwise… I do not think the general public understands how hard it is to be going in there with a lot of empathy and sympathy, and hugs, and at the same time you're worried about getting shot. Ambushed. It is a tough role and I think our officers do a great job with it… and it is hard to identify the underlying issues. If it were easy to identify, it would already be fixed, this [SDI] data can then help that way. [P10]

SDI is an effort towards a proactive policing mindset that helps shape police agents' crime conceptions to try to proactively address crime's underlying issues. Data-driven practices may alter the way police work within society. Data, as part of the proactive mindset, can foster collaborations between various public safety agencies (e.g., emergency medical services, fire departments) to depict a more inclusive understanding of the city's problems that utilize their resources. With this envisioned change, my participants grappled with the need to balance the dangers of the job. Within MMPD, there is resistance towards a proactive orientation by explaining that policing organizations are not equipped to handle solving underlying social issues:

> …with our resources we're just not designed to handle...we're not public prevention. We [police officers] see the end result. They're [social services] looking at the disease. We can just see and respond to the symptoms." [P4]

Some police understand their work as reactionary responses to emerging issues and position other social services as better suited for treating crime's underlying problems. The conversation within MMPD about these two orientations is in part shaping the future of their police work and identity. As MMPD commits more to being data-driven, police workers grapple with how data and technology can and cannot assuage social criticisms, while also inducing reflection on their sociotechnical visions for their work.

Precarious and Inactionable Insights

Data-driven policing practices are motivated, in part, by police's desire to make decisions that are arguably more objective, impartial, and can be rationalized through data. While using collected data can lead to empirical data-driven decisions, there is a risk that data-driven insights may still lead to erroneous insights and faulty judgments. In this section, I discuss how various kinds of policing agents, from on-the-ground officers to data analysts, voice their concerns about how data-driven insights may lead to faulty decisions. Disparities in data-driven decision making and insights emerged for my participants in two key ways: faulty inferences from and unintended consequences of being data-driven, and unfeasibility of data-driven decisions. I highlight how data literacy issues exacerbate such concerns.

My data suggests that MMPD grapples with understanding the impact of their efforts through data. MMPD defined impact as changes to crime patterns, community tips, number of people involved in police-community efforts, and other police-community interactions. For example, a participant from the Information and Intelligence Center demonstrates the struggle in correlating improvements in the community with their data-driven effort:

> …it is hard to prove…. the Chief is using the data and focusing the efforts … I think it is having an impact, but how do you prove that correlation? [P10]

The challenge of faulty inferences and unintended consequences emerged for MMPD officers through from data-driven decision making. As mentioned in the Research Context, each MMPD district participates in CompStat meetings as a data-driven accountability mechanism. As MMPD's police agents use data for decision making, disparities in their empirical judgment emerge. For example, an officer demonstrated how resource allocation based solely on crime statistics data becomes problematic:

> Here, in southwest district, when our crime rate was really low, then [the commanders] would adjust the manpower numbers and say, 'Well, we usually like to have 30 officers work in that area. For the last six months, crime stats show that the crime is down 50%. Why don't we bust those number of officers down…let's put them over here where crime is rampant.' That was not a good idea to do. Unfortunately, [with] municipalities, that is the way they work. What you always see is those numbers start to come back up over there where you just pulled those officers… but those crime stats lead to who gets what resources. [P4]

Data analysis can be insufficient in representing the on-the-ground reality. Other examples about faulty, data-driven inferences include targeting officers' patrol, which failed to incorporate the officers' lived experience and knowledge about the area, and evaluating officers based on the number of arrests and tickets, which is not representative of police impact in the community towards reducing crime. As the commanders make sense of historical data to justify resource allocation, they stand to make false inferences about how criminal patterns will and will not be impacted by changing resource allocation, policies, and practices.

Similarly, MMPD is also navigating how data use creates unintended consequences from their data-driven crime interventions. For instance, one officer explained a data-driven initiative he undertook to reduce burglaries in a certain area during summer:

> …we would tell our officers we want you to work in these areas in this particular timeframe for burglaries. Well, burglaries went down for the days we had targeted, but what was interesting is that [the new patrol patterns] displaced [the crime]. Because on our peak days, officers are out there, [and] they're more visible in these areas in these particular timeframes…peak time for burglaries is between eight and noon on Wednesday. So, what you see is less burglaries on Wednesday between eight and noon. But now, you see more burglaries on Sundays between eight and noon. So, we go back and try something else, think about why [the burglaries are] happening in the first place. [P2]

As unplanned consequences emerge, MMPD officers figure out how to adjust their approaches for mitigating unwanted outcomes. As police determine how they impact crime when they change city conditions, using data to identify crime may actually be insufficient in reducing crime. While it would be impossible to think of every consequence police decisions can have in the community, this disparity does illuminate quantitative data's insufficiency in representing reality that future efforts can be based off of.

As police navigate the infancy of their data-driven efforts, they have concerns about data literacy, defined as the ability to read, create, and communicate data as information. These concerns range from the officers on-the-ground collecting data to commanders using data for resource allocations and strategies.

> There is a data disconnect. [Commanders] see it, but they do not understand it… They really do not have a background in statistical information. [For example,] one of the commanders was presenting the SDI data for their district. And we were looking at one of the hotspots. And there was a criminal homicide score of 10. And he presented that score of 10 as 10 individual criminal homicides in that little 250-foot x 250-foot area, which actually two criminal homicides happened because each one is a score of five. [P15]

Police decision-makers do not always have a firm grasp of what the data represents when making judgments and how those decisions impact communities. Additionally, without a top-down push towards data literacy, the efforts to be data-driven remain half-hearted.

A lack of resources for data-driven efforts, including manpower, technology, and data expertise, makes implementing data-driven decision unfeasible. Here, a patrol officer reflects on how insights about vehicular accidents in the city are futile because their resources for action are limited:

> You cannot really predict the certain mile area of where a car crash will occur. We laughed because we were being told that we were going to have more directed patrols [meaning that they would get precise traffic predictions]. We know [the data] is there, but there isn't really anything we can do with it because we do not have the manpower. You can keep telling me all along that we are going to have crashes here on the interstate. But if there is only two of us working, we cannot be proactive. We are more reactive than proactive. [P5]

Proactive, data-driven efforts do not impact the officers on-the-ground because organizations still struggle with insufficient workforce. Similarly, the lack of technological and criminology expertise within the organization makes implementing certain data-driven insights unlikely. For example, participants echoed frustration with the top management's commitment to the data-driven efforts as he explained the lack of diversified expertise for data analysis:

> Sometimes I do not see the action [from the top management]. When I was first hired, we were supposed to build a staff of like ten analysts that were from all different walks of the trade. Instead of just having a GIS guy and a couple of legacy analysts, getting a couple well-versed criminologists involved whether that is from the academic side or working in a police force side. I would just like to see a holistic approach to staffing data-driven ideas… the leadership needs to make the resources available instead of just talking about it. [P15]

Without management's commitment and specialized experts in data science and criminology, their data-driven efforts remain constrained. As MMPD tries to incorporate data into their decision-making, the difficulties in implementing insights require improved data literacy, and technological and expertise resources.

Police Metis Concerns

While participants indicated data helps justify and guide police actions in the face of growing social criticisms, concerns about the balance between the role of data and experiential, practical knowledge in supporting police work and accountability emerged. In this section, I demonstrate how our participants experience the epistemological bias in being data-driven, and highlight the ways in which quantitative data is insufficient for supporting police work, determining long-term impact on communities, and bolstering their accountability efforts.

Police work entails a deep level of qualitative knowledge, referred to here as metis, which includes experiential knowledge, discretion, expert judgment, gut instinct, and biases [Goldstein, 1960; Bittner, 1967; Livingston, 1997; Scott, 1998]. For example, police judgment is useful when deciding to issue a warning instead of ticket, in building relationships and trust in various communities, when dealing with mentally ill individuals, or when dealing with interpersonal conflicts and minor misconducts. My participants raised concerns about overreliance on data-based insights that neglected police metis. This becomes problematic when data-driven efforts epistemologically favor metrics that can be easily quantified over officer's lived experiences. Examples of qualitative knowledge in police include the kinds of conversations and relationships police have with community members, the types of individuals police engage with, the times and places to patrol accounting for on-the-ground experiences, or how to frame police documentation about community interactions.

Several participants pushed back on data-driven insights and their ability to sufficiently inform police work. Because experiential and practical knowledge is essential to the service and practices police engage in, MMPD police officers emphasized the importance of qualitative knowledge in their work (i.e., experience, knowledge about the community, relationships, and gut instinct). MMPD's police force grapples with understanding the changing role of police metis when technology and data do and do not support their work. One participant spoke about how he thinks about the role of data in relation to the "human element" of their job. Data supplements work but the experiential and practical component of police work holds more value because of the way experience lends itself to their human-centered work:

> I work almost totally on the data side right now, but I will be the first to say that the human side of [policing] is the most important […] The data side of it is extremely important too. It can help show the crime that was committed. But beyond that there's still the whole investigative side that has to be done where [police] have to engage in that human element to get confessions and intel that data can't provide. There is nothing better than a human being doing what they do, especially police officers, out on the street, face-to-face contact, talking to people. [P8]

Here, P8 refers to engaging with the "human element" as a skill used to derive information out of individuals and communities they work with. While most of my participants acknowledged data's potential to support police work, several participants were also quick to temper the role of data in comparison to the "human element" of their jobs. Similarly, a cybercrimes analyst echoed how important it is for police officers to find the middle ground between their own judgment and data:

> Well, we cannot take all threats and all incidents and then put them in one predicting police formula. We need worker bees on the ground that take these incidents and analyze them from not only a data perspective, but a human intel and intention perspective as well. You cannot just police from a computer… data doesn't go take the runs or interact with community members, solve their problems, get confessions. So you can't have administrators making decisions on data, [if] they aren't in the field with that human intel, (being data-driven) is not going to happen. [P7]

Similarly, P7 points out that police work cannot be heavily based on data because of the role of human interaction in determining context and "intention." The human element, to my participants, is important in all stages of policing—from taking runs to investigating crimes to administrative decision-making, and data-driven strategies bring concerns regarding over-reliance on technology. While quantitative data can help police create seemingly objective pictures of their actions, the limitations of quantitative data in supporting police decision-making and work continue to emerge.

Another limitation of quantitative data is how the use of technology and data impacts police capacities, skills, and communities. From these participants' perspective, being data-driven may over-rely on technology and data in ways detrimental to police work and community. For instance, a participant, when speaking about the generational gap in technology usage within the organization, voiced his concern about how technology's prevalence shapes and interferes with younger officer's capacities:

> …when it comes to the older officers, face to face, boots on their ground, real old-fashioned investigation will always, always trump technology... Data is just [a] supplement. In the end, a detective still has to exhaust that information and go out and interview or do something and find out if this is the right person. What I actually worry more about is the younger officers who start relying so much on technology. [Younger police officers have not learned] how to do that people side of the work. … Sometimes [I have to] pull them out of the computer and say, 'you've got a human being sitting down at the entrance, who has information, go talk to him.' Or some of the biggest carriers of information for police [are] your drug dealers and your prostitutes. The younger officers are not really comfortable street wise, with the social dynamic of talking to people. [P8]

While a few participants echoed concerns about how data-driven strategies impact police work, P8 indicated that the "people side" of their job is where police impact their communities; quantitative data is insufficient for capturing impact on community relationships as the social skills required for policing are at risk for these participants.

Lastly, quantitative data also falls short in the promises for providing increased accountability for police actions and community relationships. While most of my participants talked of data and technology as largely objective and neutral, one participant directly discussed the biases embedded in data work as well as police work. The participant explained how data, curated to be reported, in community meetings can be shaped into certain narratives:

> Biggest challenge with data is accuracy. Here's what's bad about data: if you want the numbers to look higher than what they are, then you're selective about what data you pull. You know, on the local [community] level, the numbers are going to say whatever you pull and have them say. [P4]

For this participant, a challenge he sees in quantitative data use in policing is about how data is politicized. Data is shaped based on who is collecting, analyzing, curating and presenting the data and how. Data use in policing for accountability raises issues of how data is used to shape narratives and provide rationality from certain perspectives. While quantitative policing data does provide certain lenses into police action, police's data-driven strategies are not designed for improved accountability to the community.

Big data supports data collection and interpretation of phenomena that can be easily quantified, which means it may leave out critical policing knowledge not well suited for quantification. Through the conversations surrounding the role of data and police metis, MMPD officers question quantitative data's role and legitimization. As the police use technology and data to portray their work practices as objective and neutral, they also worry how over-reliance on data devalues their metis and consequently, impacts police outcomes, specifically, the police skills and practices fundamental to working with human-beings.

5.3 Discussion

Data and Police Judgement

As concerns about police judgment grow in terms of police's disproportionate and negative impact on marginalized communities, data and technology stand to drive decision-making and police actions [Morgan et. al. 1997; Manning, 2005; Ericson & Haggerty, 1997]. My results indicate that polices' data use alone is not sufficient to support police action and judgment because quantitative data provides rationalities that hinders police's ability to understand marginalized communities' policing criticisms and thus, hampers efforts towards accountability. My research emphasizes the importance of understanding in what ways and how data can and cannot support police work.

Now, I discuss how data *could* be used more constructively to influence police judgment by facilitating collaborations between the police and community to more robustly inform police's preconceptions. First, quantitative data can be helpful in providing aggregated views of social issues that perpetuate criminal patterns. One of the frustrations my participants expressed was that while they understand how police actions impact communities negatively, directly addressing crimes' comorbid issues of poverty, mental illness, and drug addiction are beyond typical police jurisdiction and training. While other scholars argue that police play a role in exacerbating these issues [Bitner, 1967; Jacobs & Britt, 1979], the role of data in confronting police criticisms could help expand the issues police focus on and provide police with alternative ways of addressing issues. For example, instead of focusing on enforcing the law, through collaborations with social workers, police could be equipped to provide resources (*e.g.*, food access, occupational trainings, or job opportunities) to meet community needs. Second, data can temper police judgement through demonstrating police biases. Data can either confirm or counter their biases, depending on how data collection and analysis occurs. For instance, data could be used to demonstrate police bias driven by preconceptions of poverty and/or race (*i.e.* the kinds of cars pulled over, the kinds of people targeted, and the kinds of leeway given to crime in higher social status communities). Third, quantitative data can support evidence-driven approaches towards community interventions. Because police agencies (like most public services) operate under significant resource constraints, proactive, community-building efforts, guided by big data, can address community needs, and improve access to resources (*i.e.* youth programs to provide positive role-models, educational meetings for immigrant communities, drug awareness campaigns, or social work resources).

5.4 Conclusion

For the participants at MMPD, challenges of adopting data-driven strategies in policing are experienced as they negotiate the sociotechnical visions for their work, some of which were explicitly meant to confront social criticisms of policing. In this study, I have made the following contributions: identified three key challenges in adopting of data-driven policing; identified how data can support and constrain police work and accountability; provided design implications for data use in police accountability.

Chapter 6: Re-Orienting Data-Driven Mythologies

6.1 Research Context

When this research around data-driven policing began, as indicated in Chapter 5, the Obama administration had already introduced optional data-driven initiatives for the police to participate in for improving community relationships and accountability. However, by the time the second study began with a diverse group of stakeholders in policing, the U.S. has been under the Trump administration. Subsequently, the fate of the Obama-era data-driven policing programs remains unclear, since this information is now archived online[8]. This cultural and political context is important to understand and situate this research because of how this administration is influencing the conversations around political issues of racism, marginalization, police accountability, immigration, etc. President Trump and his administration have been stripping away civil liberties and protections across several departments of the government (police accountability, education, student loan protections, etc.) [Minhaj, 2019]. The political rhetoric, since the 2016 election, has also demonstrated extreme polarization, with critics blaming the emboldened, discriminatory narratives that the President pushes, citing 17% rise in hate crimes within Trump's first year of presidency [Rogers et. al. 2019]. Rogers notes that this polarization is not solely because of the president, but also other factors in society. For instance, social media platforms are also under fire for how their algorithms induce echo chambers, where people are exposed to opinions they agree with [Garimella et. al., 2018], as opposed to accessing impartial news around important social issues.

[8] https://obamawhitehouse.archives.gov/blog/2015/05/18/launching-police-data-initiative

Similarly, the concept of "fake news" (which emerges in my research often) has been adopted by President Trump to discredit often legitimate and critical information [Lazer et. al. 2018]. The "fake news" claims are intertwined with the epidemic of misinformation that is being spread through social media [Satariano, 2019]. However, on another hand, some consider the election of Donald Trump a turning point in perhaps spurring more citizens to participate in democracy (*i.e.,* record setting turnout for 2018 midterm elections) [Domonoske, 2018]. The levels of participation indicate that the existing and emerging political issues are important for constituents to influence with their votes. Beyond elections, there continues to be a wide space of grassroots activism that continues to do work around politically contentious topics like police brutality, with increasing attention being paid to how the rhetoric espoused by the President is impacting marginalized communities [Levin, 2018]. In order to understand the results that are presented in the next section, it is important to situate the research context politically and culturally here.

6.2 Results

In the following section, I have identified three ways that stakeholders perceive data use as an extension of police's power in the quest for accountability. Who has power in society manifests in several ways, based on which groups of people get to influence access to and decisions around creating and legitimizing knowledge [Foucault, 1980]. My participants raised concerns around how they see power disparities impact police-community relationships from varying perspectives and how data stands to perpetuate those disparities, while hindering police accountability. In what follows, I outline my participants' concerns around police's power manifesting in how knowledge is created through data, how resources are supported through data, and finally, how police control communities through data.

Data Use for Knowledge

In this research, power disparities around data-driven police accountability primarily emerged around which groups of people get to create, legitimize, and use data as actionable knowledge and what those groups' values, interests, and biases mean for the space for action in different communities. Here, knowledge is defined as representations through data that are legitimized as facts, information, and narratives around police accountability. While data's politicized nature is widely acknowledged by academics, data-driven technologies and practices are marketed and adopted with the promises of establishing ground truth, supporting decision-making, and achieving accountability. However, in this research, it was evident that most stakeholders were aware that police's data could not be unbiased, primarily because how political police accountability is in our society. Despite the level of trust the participants had in the police, almost all participants expressed mistrust in data's ability to present the truth around a contentious topic like police accountability (at least in its current form, in the current political climate).

Manifestation of Power in Data Use

Organizers and community members, who were critical of the police, were quick to point out how the creation and use of data manifest power imbalances, such as not including (and legitimizing) marginalized communities' experiences and reinforcing existing, systemic inequalities and biases. For example, an organizer here points out how data produced by the police, about the police, could not produce sufficient accountability between police and the communities:

> Police accountability, with how things stand in our country, is impossible. Maybe forced accountability is an option? When one side has a gun pointed at the other, there is no chance of them 'serving' us. So it doesn't really matter to me what data the police can show me because they're making that shit up themselves, it's not like they follow any standards set up by anyone. They are creating and showing us information about their own selves, holding their own selves accountable, how does that work? We don't need to see more crime data and more rationalization for them policing black and brown bodies. That's one sided… is [the data] really reflecting what black and brown communities are going through at the hands of these officers? [P8]

While P8's frustration is at large with the system of laws and policies that allow police to shape accountability practices, her comments around the role of data in trying to produce accountability sheds light on the facts that the data presented by the police is often obscure. The constraints around how the data is collected, curated, and presented are not clear in the first place. In my previous study, the police participants also indicated that data collection and creation is not an exact science and is often an arbitrary process district to district, marred with entry errors and legacy systems. The organizer also points out that this data is imbalanced in whose experience it depicts. Not only does this data use fail to represent the reality for POC communities, it also does not show the information around police actions in the communities to be able to hold the police accountable overall. Similarly, another organizer reflected on how police's data use and policies are exclusive of the experiences of the people most impacted by the police:

> I am not here to understand policies that result in dead community members. I am here to rewrite policies that are inclusive, that listen to those who are the ones most affected and then take direction from them on how we proceed. If you continually have people at the table that do not have direct experience with the problem you are trying to fix then your solution will never be about solving problems. Your solution is about making yourself feel better. We have to get to a place where we are humble enough to understand that listening solves many more problems than deciding to do things TO people. [P15]

Without the input of marginalized groups, police's data-driven accountability efforts remain dubious for several of the organizing participants. The spaces where decisions about what data to create, how to use it, and how to present it, are imbued with inherent power. Not only are police officers the ones in charge of the data creation and use, the data they are *able* to collect is inherently incomplete. As another organizer pointed out, community policing applications and forums stand to exacerbate how POC communities are policed:

> Something like NextDoor can be helpful but …when you have a NextDoor community that is mostly white people and an actual community that is mixed demographically - racially, economically, and everything else, but it's seen as a solution to the problem of undesirables and it's leveraged in a way that is based on fear of something happening versus something actually happening. So that's all information going to the police and it is then used to disrupt actual communities because now white people have another [virtual] space to be scared of 'others'. [P3]

P3 indicates that the groups of people that trust *and* choose to engage with the police and report crime tips does not necessarily depict the on-the-ground make-up of a neighborhood community; this organizer raises concerns around how applications like NextDoor can exacerbate policing of minority communities. While the organizer indicates that the "fear" of "others" drives who gets policed and for whom, community data sources are limited in who the police can collect information from.

Several of the participants indicated limited optimism that a technological solution could help mitigate the power inequalities between police and marginalized communities because of how they perceived technology (through both design and practice) to embody the underlying assumptions and biases of the groups with power. For the organizing and critical community participants, it was evident that if the data police collect and use is to represent a more grounded truth, the power imbalances around which groups of people and communities get to influence data practices must be addressed.

It is also important to note that some of my participants saw POC communities' hesitation and refusal to engaging with MMPD as a symptom of data being transactional for accountability. The police participants expressed frustration with the lack of engagement they receive from the community when they organize trainings and events to inform and disseminate data to the community ("if we organize these academies and public safety trainings for a hundred people, and then we keep having only 20-30 people sign up year after year, then how much is that our fault? People are not invested in law enforcement. They don't recognize it as their own." [P9]). However, through observations, it was clear that attending informational events that the police organize, outside of community meetings, often meant filling out an application for a background check, which includes giving up personal information like Social Security Number, birth date, living history, work information, etc. A couple of my participants saw these events as transactional and limiting the inclusion of different groups of people, who would either not apply because of discomfort around not sharing that personal information or simply because they might not make the cut from the background check. While the police's intention to share information with the community and be more transparent is evident in the events and information they *do* offer, this knowledge is still guarded and exchanged for information that might be uncomfortable for some people to transact with the police.

Additionally, there was a stark difference in how the different stakeholders experienced access to police and information around the police in this research. Community participants who were part of local organizations like Crime Watch indicated (boasted about) the level of access they have experienced with MMPD. For instance, one Crime Watch community member here shared an instance where he received complaints about a stray dog from his neighbors and how he was able to decide what to do in that situation (which happened to be to shoot the dog because he was "aggressive" and "barking") because of his connections to the local police officers:

> with being involved like this with the community and [the police], I've made a lot of connections… When I had the problem with that free dog in our neighborhood… I called [a police officer] right there, and he picked up right away, completely helpful. He was like, if you need to put down that dog, you have to. If it's a threat to the neighborhood, do what needs to be done. Came by afterwards to help wrap it up… [P7]

While P7's example of his access to the local police demonstrates the differences in experiences of different community members around police responsiveness, the local activists and more critical community participants indicated that access and utility of police's data was inadequate. For example, here an organizer points out that information about police complaint board and merit board meetings, which are the main spaces for accountability between the police and community, was obscured and inaccessible:

> I mean, they make it so hard to even be in the same space where police accountability is supposed to exist. For example, the merit board review meetings, where officers who have killed black men in this city are supposed to be fired, are done at the most inconvenient times, in the middle of the week, middle of the day. Which, okay, that's fine, but to even find out the schedule, you have to drive to downtown, pay for parking, go up to the top floor, go through this maze to a tiny, little bulletin board. Like its 2018. If you really wanted the public to be there and hold you accountable and engage, you really can't think of other ways of sharing this information to us? [P2]

Almost all community members, regardless for the extent of their support or disregard for the police, indicated that MMPD should work to make relevant information more easily accessible. The transaction of data and information manifests the underlying power police hold in terms of how easily and openly data and what kinds of data is shared with which groups of people.

Manifestation of Mistrust in Data Use

Most of my participants experienced data creation and use as an extension of powers also through the mistrust around how data is inherently framed into narratives based on the embedded biases and interests of groups of people who have the power to legitimize the data. However, while community members and organizers expressed limited trust in police's data use for rationalizing decision-making, the police on the other hand, also ascribed to significant doubt in the anti-police narratives and information they see online and in the media. For instance, this detective sees social media as an amplifier of police issues that he does not think are prevalent on the ground to the extent being portrayed:

> Social media's anonymity doesn't allow repercussions of what people say. Law enforcement has always had a healthy protest environment and counter culture. It's always going to have that opposition, if they didn't we wouldn't be in business. Officers have to work on overcoming the opposition on an individual, case by case basis. And now this is susceptible to influence from outer countries, as we are seeing. Those things are also impacting trust and how people think of our accountability I think. But if we really look at the scale of our interactions and the negative ones in reality, it would blow people away. We can do that, look at all the interactions police have against the negative ones, like shootings, force, what have you, that data is available. Are there true cops that are not fit to be on streets? Yes, absolutely. Better supervision, yes. We are always going to be organizational problems, but people don't know the scale. I don't trust the data we see online or in the media. [P4]

Here, the officer caveats his doubt in the public scrutiny around policing with the acknowledgement that there are inept police officers in the force, who abuse their power, but overall, he indicates that data could help them prove that police criticisms are not as widespread as social media and media might portray them to be.

Through these conversations with the various stakeholders, the limitation of trust in data emerged as being understandably interlinked with popular narratives around political bias in the information presented. For instance, almost all participants pointed towards the limitation of data use in supporting trust, producing unbiased accounts of truth, and ultimately, in influencing people's viewpoints. One police supporter participant, when talking about moments where MMPD has been under scrutiny, revealed that information and data around police brutality are not always important to her in maintaining trust:

> …when [an officer] hit the motorcyclist and killed, he had been drinking, there were screams of a cover up and people demoted. I didn't really question it, but that was a time I could have, I didn't let my trust waiver in them. We have to protect our own, as a group… [The chief] tried to keep things from getting out of hand. Every time something is repeated, it is different. And sometimes that looks like cover-up. Does that make sense? People were screaming for all the details, it was like a lynch mob. My reality is my perception, that's just the way it is. [P6]

Here, the participant chose not the "let [her] trust waiver" in the police when there was an incident around police accountability, despite the information that was and was not made available. Similarly, this sentiment was also echoed when another participant expressed how police accountability was a political issue and trust in information and data presented was influenced by political biases:

> I quit getting newspapers because it's too left wing. Of course, we get some data from the police, but a lot of data is reported by the media, right? Most of that is opinion. First time in American history that fake news has become a common household word. If I didn't see it or hear it, I don't believe it. [P7]

Around a topic that is so deeply politically and historically contentious, almost all of the participants indicated mistrust in the information and data that is available in the mainstream media platforms. On the other end of the political spectrum, most of the organizing participants also did not believe that accountability through data-driven policing was possible. One participant here talks about how bias against the police is important in her work as an activist around police:

> I am never going to apologize for giving a black person who says they've been brutalized the benefit of the doubt. I would much rather be wrong about that than dismissing someone who says they've experienced violence... Because also, I'm not going to be sorry for ever doubting the police. Being skeptical of them cause I know police and nothing can change those things. That's a bias, but you can bet it is earned. [P13]

For P13, information around police brutality is not as credible as a person's experience, which is why she expressed how her mistrust of the police guides her work and relationships around police accountability. This raises questions around whose data and experiences are captured and legitimized ultimately in police data, which is important because as systemic biases have been recognized and critiqued, expecting trust in data from the same system becomes unrealistic for these stakeholders. Moreover, critical participants emphasized that data can be another way for police to frame narratives around societal issues and POC communities. For instance, here a community member expressed her frustration around how the data around policing and from police carried racial biases:

> Drug support is different for white people, while black people are arrested. If you are certain color, you are perpetuating the problem, if you are white, you are the victim of the problem. [That's what is] morally unsound where not everyone is being viewed the same, as a noble, valuable being. It is how we are educated, it's our media, the way things are presented, whatever you wanna say. [P10]

This community member's critique of how police use data sheds light on how stakeholders' concerns emerge around data shaping the spaces for police actions: where certain communities are criminalized for social issues, others receive empathetic support and resources. While it was clear that almost all participants carried varying levels of distrust in data sources around policing, several of the participants indicated that more comprehensive data focused on police's activities and their interactions could possibly counteract how data is framed into narratives.

> They're always going to heavily police our communities, they'll give you this data and that data, but we don't get data about the police and what they do. The FBI has reports about crime data that police collect for them, but not any standards that force police to tell us how many use of force cases they were involved in, how many unarmed shootings, how many cops have a history of abuse. We are just starting to get those information from other activists and organizations, websites where regular folks, citizens, are doing the work to hold police accountable with data that can actually paint that picture [of police actions]. [P15]

If data is to support police accountability, stakeholders expressed wanting to see police present data on their own activities, as opposed to just crime data. As evident by these empirical results, stakeholders' concerns were around how data can be an extension of police's power in what societal issues are criminalized, which communities are policed, what narratives are legitimized, and how they rationalize police's actions.

Data Use for Resources

While several participants were critical of how data use shapes narratives and information around the police, police's data use also raised concerns around how decisions around resources are rationalized and allocated within policing organizations. Resources here refer to expertise, technology, materials, information sources, funding, and other assets to support a police organization's work and practices; resources also include the services the police provide to the community.

For most of the participants (community, organizers, and police), there was a succinct awareness that it is impossible for the police to collect *all* data around the city's police and crime issues. Whether the participants attributed this to the fact that not all crimes are reported by community members or that police do not patrol all communities with the same suspicion and leniency, several participants saw technology and data use as likely exacerbating existing biases and even leading to over policing of POC communities. For example, P3 further commented about applications like NextDoor that can present a skewed reality based on which groups of people participate in those spaces and create data for the police to utilize:

> … the thing that is still coming to the light like with activists working in Chicago and New York and we are trying to understand is how these datasets, tips and community reports, and police data are used to put police officers where they patrol, how the brass use data to justify the surveillance and policing of marginalized [communities]… [P3]

Here, this organizer echoed other critical participants about the obscurity around data-driven decision making within police and how technology and data influence the interactions between police and different communities. If certain communities are not consciously included in social and technical platforms, that data source can stand to perpetuate the existing power disparities between police and POC communities. For instance, several community participants from crime watch groups bragged about their access to the police when they have had a concern or an incident occur. As demonstrated above, P7 shared an example of a neighborhood disturbance and how his access to the police enabled him to get help faster than he knows a regular citizen would:

> it's not about who you know, but who knows you. That's what gets [an officer] to text you back immediately, or get all those potholes fixed. [P7]

Inversely, several other community and activists were dubious about police's response rates to different communities. Here, a community member talked about her experiences around police responsiveness after moving from a predominantly minority neighborhood only a couple of miles away:

> Yeah, and the police respond very quickly here. So we have a young man who used to live in our former neighborhood, it's a 10 minute walk away,

> but when we moved to this mostly white neighborhood, we were surprised about how the police and emergency services respond in like 3 minutes here, but back there, whenever there was an emergency where we used to live, it could take the police or even the fire department a really long [time] to get there. So if like something goes wrong, you know that the police are going to be there, the fire fighters are going to be there, and that was coming from like a young refugee boy who has lived both in both these neighborhoods just a short distance apart. [P5]

Several other participants also echoed the same doubts around what factors influence how responsive the police is to some communities compared to others in the city. Concerns around how the police relate and respond to different communities were also intertwined with concerns around how data can stand to make access to police better and worse for different communities:

> …if only richer folks, white neighborhoods or people are tipping the police, posting and creating these leads for the police, all this community-generated data that the police always push for, it is just an extension of the surveillance of police on marginalized communities. So we don't ever really see our communities' voices and concerns in the information police have and work with. Could we? I don't know, I really don't think so, not through [the police]. [P8]

If data use shapes the space for actions, legitimizing needs, then it is important to address whose needs are vocalized through data and whose are excluded when only certain communities choose and/or are able to engage in police accountability spaces and practices in sociotechnical ways. For these participants, data use simply cycles through the same power disparities and racial biases that already exist in how different communities are surveilled and served.

A few stakeholders also expressed wanting accountability around the resources and funding police receive as well as the role data plays in acquiring those resources. These concerns, mainly from organizers and community members, revolved around the increasing militarization and surveillance state of the police. For instance, another activist discussed what MMPD's investment in a data center means for the community:

> I think there are bigger things that we don't talk about in terms of where policing is headed, right? Yeah, they have this data center now and it is promoted as you know, making us all safer now as a city, but we know they are using that data to justify asking for more weaponry, more militarization as well, that's happening all over the country. If we, the people the police are supposed to be serving, don't trust and cannot even really see this information in the first place, how is it okay for them to use it to get funding? [P15]

This participant points out that while data use can support rationalization of police needs and agendas, being data-driven can be to the detriment of the marginalized communities, especially considering that the data police produce is not guided by consistent federal or state standards still. Without standardized data collection practices and procedures, another participant was doubtful that such funding requests truly depicted community needs. On the other hand, during my observations with police officers, a few of them did view data as increasingly more essential in "proving [their] case" [P4] for investments from governmental funding agencies. This was especially important to them because of the manpower and technological expertise constraints they operate under. However, one officer, during a ride-along, also expressed confidence in being able to acquire funding more easily for weaponry and technology under the Trump administration. Ultimately, for several of these stakeholders, data use supports and manifests political agendas that are already in place within police accountability issues, instead of a way to truly depict community needs and provide rigorous accountability around police actions.

Several other participants, including a couple of police officers, expressed hope for police's data use to identify and legitimize underpinning societal issues related to crime, such as mental illness, poverty, drug abuse, etc. While community and organizer participants shared this view, one police officer in particular defensively voiced his frustration with the larger political infrastructure that does not provide the necessary services for at-risk community members, placing additional burden on police agencies:

> …it is not our fault that from the top, legislation, public services and funding for mental illness related issues and, you know, public safety institutions, gun control, you name it, has been gutted over the past decades. So now there are calls for us to be trained in this and that, but there are no laws mandating those trainings as a profession, the expectations are not shared between the police and the community. So yes, I believe we [the police] have always been held accountable. Does that mean it's to the same level that some people expect? No, probably not, that's not going to happen unless actual legislation comes in, so all the public and media can yell at us, sure... [P9]

None of the police participants acknowledged that police accountability in MMPD had failed. The hedge in P9's statement is that they have been accountable according to the standards and laws that are set, but he also knows that the level of accountability mandated through legislation is not always up to the communities' expectations. From the police perspective, accountability is often about how our federal and state laws do and do not operationalize police behavior and training.

If police services are a resource to protect and serve communities, several of the participants indicated doubt around how policing resources are supported within the larger political system. For instance, one community member, who did not think that the police are held sufficiently accountable, did point out that police are not functioning in a "morally sound" political system:

> …[we have] laws that are criminalizing poverty, drug abuse… overall, our government, society do not provide a morally sound framework for [how police] can serve communities and how police could even be used to uplift struggling communities. [P5]

Resources in this context revolve around both the support and services that police agencies receive as well as the communities that are served by the police. My stakeholders expressed concerns around how data use impacts the disproportionate amount of resources police stand to receive that could further marginalize POC communities, as well as how police are often held responsible for lack of resources to address underpinning issues that lead to crime.

Data Use for Control

From the first two empirical sections, it is clear that stakeholders are concerned around how data is transformed into knowledge and how that knowledge then is used to rationalize police's decision-making. Subsequently, these concerns, particularly from organizers and POC community members, settled around how police use their power to profile and surveil marginalized communities, and the role data plays in supporting those spaces of police action. While access to resources is intertwined with police's capacity to control communities, control here refers to the discretion and responsibility police have to maintain social order, through force if necessary. While data can impact what narratives are legitimized and what resources police function with, for my stakeholders, data also stands to play a role in the interactions police have with community members, particularly POC. For example, almost all of the organizer participants and POC community participants shared examples of times they had experienced profiling; here an activist shared an incident where she filed a complaint against police officers over being pulled over as a suspect, even though her vehicle's description did not match the information the police had:

> I have a couple of police misconduct complaints going on with [MMPD]. Like one when they pulled me over at McDonalds and those officers knew me, but were looking for a suspect, they had information on their description and car. Of course, turns out the description of the car was so far off, totally different color and make but they all know me over at MMPD. They will always say that we are just investigating, doing our job,

> but why would you pull over a completely different car and act like it matches your description when it doesn't? I know they have all kinds of information on me. So that was just an excuse for them to come harass me as they do... [P15]

Here the organizer speculates that she was targeted in this instance because of her contentious relationship with the local police, where the police misused the suspect information they had to pull her over, which then became the basis of her complaint against them. Similarly, another activist shared her experience of being targeted and intimidated by the police:

> Like I said before, there have been several times that I have been targeted by the police, and it really makes you wonder what data they have on you and how it all flows between them. One time, the FBI came to my parent's house, looking for me. I don't live there and hadn't for years. Another time, that same year, the local police pulled me over and he greeted me with my name. It's like we know your name, we know about you, we are watching you… [P8]

While racial profiling and intimidation are prevalent in the experiences of POC and activist communities, for most of these participants, they saw data as another way of giving the police tools and opportunities to target marginalized communities. Similarly, all of the POC community participants shared several examples of precautions they take in interacting with the police, as well as examples of interactions that might have been motivated by officers' racial bias. For these participants, the role of data was questioned in how surveillance and profiling is justified and motivated around different community members:

> I think I worry for my brothers, as black men, in this society a lot more than I do for myself. I mean, my younger brother didn't tell me until years after about how much he used to get pulled over by the police when he was in high school, when he first started driving and like for nothing, cause we didn't even know at home, right? But that's all data that I think could be so helpful. Right now, it seems like the police are using data to justify their racist patrolling and policing practices, but when do we get to see that overall data, that these officers always pull over this many

> minorities and white people and how much does it actually end up in some kind of charge or unarmed shootings? What are the races of those officers? Racism is a problem in this country with the police, and those in power ignore this in everything they do. [P5]

This community member expressed that while she worries data is being used to perpetuate racial bias within policing, she would much rather see data used to paint an aggregate understanding of how race plays in a role in police-community interactions. Several of my participants indicated that unless the issues around racism are better understood and legitimized, police stand to continue racial bias through their data-driven policing.

Inversely, while none of the police participants indicated that police accountability had ever failed within MMPD in their experience, one officer did acknowledge the inherent power that comes with being a part of a policing organization.

> …what it really boils down to is that the community has no choice but to trust the police. Everyone calls us, and the ones who are the loudest against us are the ones, funnily enough, calling us even more… look at that data, right? Even if they hate us, everyone still calls us for everything. Most of the times for things we can't do much about anyway. [P12]

Here, P12 acknowledges that communities do not often have another option or avenue for resolving issues, despite how they might feel about the police or how much the police could help them ultimately. Similarly, another officer expressed doubt around the legitimacy of activist organizations as he raised concerns around who funds BLM activists:

> There has always been a healthy counter culture against policing. BLM isn't new, but what we see as officers is that these issues are politicized right? So you have to look at the funding these activists get to disrupt the communities, Soros for example. That's part of what we do on social media when we monitor and observe certain groups. Have there been injustices to them? I'm sure. But these protests and things are part of larger agendas. [P4]

Here the officer indicates that police use social media data to monitor local activist organizations because of how he thinks such organizations can be used to advance anti-police sentiments in society. Police represent an infrastructure of dealing with conflicts and safety-related issues, which is both about police's responsibility and power in the community. Data's role in maintaining and exacerbating these power disparities stands to be problematic.

6.3 Conclusion

In this section, I have identified three ways that my stakeholders experience data use as an extension of police's power, as well as outlined how those stakeholder perceptions hinder police accountability. The legitimacy of data produced by the police is often dubious for several of my participants because they see data-driven police accountability as biased and exclusionary as the criminal justice system it is situated in. Similarly, trust in police's data is also impacted by the larger cultural phenomenon of misinformation.

Chapter 7: Design Implications

This case study of data politics in human-services highlights larger societal issues of power, unequal citizenship, systemic biases, and governmental accountability. My dissertation contributes at the intersection of data use, accountability, and power disparities. This research demonstrates several challenges and limitations of data-driven technology and practices, such as how values associated with a “culture of data” lack social and human context; epistemological biases towards certain forms of data; frictions between data and domain expertise; political mistrust of representations of reality; and how data use often perpetuates power disparities and systemic inequalities. Subsequently, I aim to reorient the data mythologies that I have identified through this work to mitigate the power disparities that hinder data-driven accountability. Because data’s embedded values and biases are inevitable, I present design guidelines around how data-driven values of objectivity, trust, and transparency can be reoriented through pragmatic HCI guidelines to support more robust accountability. However, before I dive into the design recommendations, in what follows, I present arguments regarding the political and ideological context underpinning the work around police accountability and how these contexts are intertwined with my design guidelines around data use.

The key stakeholders in this research adopt sometimes incompatible approaches (abolition vs. reform) around the issues of police brutality and accountability. Abolitionists fundamentally question the legitimacy of police organizations, advocating for alternative ways for public safety and conflict resolution than the current criminal justice system [Purnell, 2017]. Police abolitionists’ rhetoric is rooted in the history and identity of unequal citizenship, which refers to minority groups who have not been given equal voice for democratic participation [Collins, 2012]. On the other hand, reformists are working, to varying degrees, within the criminal justice system to make incremental changes to how police and courts disproportionately impact marginalized communities. In other words, reformists focus on how to cultivate pathways to equal citizenship within the system [Meyerson, 2004]. Both approaches, while in several ways can be irreconcilable (*i.e.,* my abolitionist participants refused to engage at all with the police for demanding change, but rather invested their time and efforts into creating new ways of community organization), hinge on the shared mistrust towards the police. With this research, I argue that both these approaches are necessary and intertwined in achieving a more equitable futures around policing and marginalized communities. Subsequently, my dissertation advocates for how HCI can support both approaches through being data-driven.

In this research, while activist and police approaches and attitudes towards public safety are different, ultimately, they do share the goal of community safety and equality. This means that there are overlaps in how technology and data use can help achieve accountability around the issues of crime and misuse of power. While politics and systemic inequalities are responsible for creating social issues, such as police accountability, I content that technology use can and must play an important role in dismantling existing biases and inequalities instead of exacerbating and perpetuating them. The biases and agendas embedded in technology need to be evaluated and mitigated as data-driven technology and practices become more prevalent in our society. In this discussion section, I unpack more concretely how technology and data use can better support police accountability.

7.1 Designing for Strong Objectivity

Research around data use questions its objectivity and relativism—what are valid ways of creating and legitimizing different kinds of knowledge? As Sandra Harding points out, to choose between objectivity and relativism is inherently limiting to our ability to understand "nature and social relations" [Harding, 1995 p. 332]. She positions the concept of "strong objectivity" against scientific objectivity, which has often been portrayed as neutral. However, scientific objectivity is a widely contested concept. First, objectivity has been, historically, attributed to certain groups of people, whereas other groups have been characterized as less capable of making impartial and objective judgements [Harding, 1995]. Second, objectivity relates to how methods of knowing are considered "fair" or more objective in producing knowledge claims (*e.g.,* quantitative vs. qualitative methods). Finally, scientific objectivity is contested because of how certain groups of experts include and exclude "members of different classes, races and/or genders" [Harding, 1995 p. 333]. Similarly, Proctor (1991) characterizes neutrality, which is a requirement of objectivity, as a "myth, mask, shield and sword" because of how it can shape problem spaces for action. Feminist and post-colonial scholars have demonstrated that the concepts of objectivity and neutrality are fundamentally flawed and thus lead to narrow and limited ways of knowing about the world, which are often imbued with values and interests of the powerful groups. As Harding notes:

> Objectivism defends and legitimates the institutions and practices through which the distortions and their often exploitative consequences are generated. It certifies as value-neutral, normal, natural, and therefore not political at all the policies and practices through which powerful groups can gain the information and explanations that they need to advance their priorities. [Harding, 1995 p. 337]

Instead, strong objectivity, which here is defined as being inclusive of other perspectives, strengthens standards of objectivity. She claims that in order to "modernize the notion of objectivity," we must strive to "maximize objectivity." While we cannot achieve absolute objectivity, we can adopt practices that view objectivity on a spectrum from weak to strong. Strong objectivity entails including perspectives from multiple vantage points, different groups of people for more robust knowledge creation. In this dissertation, I apply the concept of strong objectivity to the rhetoric and limitations around data use and how this viewpoint can help us mitigate data's false claims of neutrality. Similarly, the concept of strong objectivity also informed the research design for the last study in terms of collecting data from multiple groups of stakeholders.

As data-driven technology and practices becomes more normalized and popular in society, it raises questions like how and what knowledge is created? Who gets to create said knowledge? Which data are legitimized? Whose experiences, interests, and biases are embodied in the data? What underpinning values and societal inequalities do data use carry?

In what follows, I outline three main ways that strong objectivity can inform HCI at the intersection of practice, design and policy around data use: participation from key stakeholders (practice), legitimization of metis and non-traditional forms of data (design), and regulation of data creation and use policies (policy).

Participation of Key Stakeholders

My results suggest that as claims of objectivity and truth are integrated into data-driven organizational cultures, HCI scholars and designers must address limitations around data use and mitigate power disparities (*i.e.,* financial, educational, legal, safety, etc.) that emerge on-the-ground. This is especially important because organizations with power, like the police, are using data to establish ground truth, which then stands to shape the kinds of support and treatment communities receive. Ultimately, this research raises concerns around how one group of people can investigate and hold themselves truly accountable, as well as how like-minded groups of people can reveal their own biases that manifest in data. As evidenced by this research, and the current cultural climate, which is characterized with significant political polarization (as pointed out in Chapter 6), people from various communities feel strongly about police accountability across the political spectrum. It is clear that just as police are heavily scrutinized, so too is the data validity that is produced by the police.

In this section, I address how community participation from key stakeholders can help maximize objectivity in the police's use of data and potentially address some data limitations in achieving police accountability in their data use. Based on how data is shaped and decontextualized, data makes certain parts of police work invisible, including work practices and stakeholders (beneficiaries and marginalized communities). The practices that are not easily quantified and digitized are relevant to account for in understanding how police shape data and impact communities. My results indicate that polices' data use alone is not sufficient to support police action and judgment because such quantitative data produced by the police provides rationalities for action that hinders their ability to understand marginalized communities' criticisms of the police and thus, hampers efforts towards accountability.

Consequently, my research shows that solutions for combating issues of data politics in police accountability cannot be effective without deeper community engagement from key stakeholders. Previous work in HCI has highlighted the role of technology in community practices around crime [Blom et. al., 2010; Erete, 2013]. For instance, Erete's work indicates that while technology can help increase social capital and efficacy in neighborhoods, technology is better suited to supplement community meetings around crime, rather than replace them as citizens participate in local, civic engagement initiatives [Erete, 2015]. Community informatics research has demonstrated that technology by itself may not increase communities' political power, but does provide mechanisms for governmental accountability as well as options for citizens to participate in local decision-making [Erete, 2013; Erete & Burrell, 2017].

In order to assist police in becoming more accountable, I recommend that all phases of data work be community-driven. Data work here refers to the planning, creation, analysis, and presentation of data as knowledge with a larger committee of key stakeholders. Such community-driven data practices could enable a more transparent way to account for police practices as well as determine how police actions are reflected in the data. This kind of collaboration would pull in stakeholders, including community members, critics, activists, and data experts, from civilian society to be part of data planning (*i.e.* data collection, analysis, and implementation). Currently, the community is presented with data as a final product of police work. However, if police want data to legitimize their work, data cannot be created and presented to the community as an end product, but rather should be shaped in collaboration with the community, particularly being inclusive of voices that are marginalized by and critical of the police. This collaborative effort could then aim to explore and implement ways that we, as a society, understand the underpinning issues of crime, recidivism, and systemic racism. Moreover, this does not mean that this committee of key stakeholders has to be open to the public since the police have to be careful of what kinds of information and decisions they can reveal. However, having a more independent committee of experts, activists, police, and community members could help regulate and evaluate the use of data, grounded in the context of the city.

However, challenges exist in involving stakeholders from minority and marginalized communities in the co-creation of police data. These include, among other things, whether the police is willingness to cooperate with these communities. Subsequently, such an initiative raises issues of power in this context. Indeed, depending on who is permitted to be involved, the data created could continue to carry certain agendas or reproduce inequalities, as it may not adequately reflect the realities of those individuals and communities who are subject to its use. Another possible criticism of this type of community-engagement panel, as described earlier in this dissertation, could be the self-selection of wealthier and well-off community members, who are looking to gain political ground locally. Likewise, data literacy is an issue that could constrain community members from participating as well.

While the components of participation from key stakeholders requires further defining, my research suggests that it could provide a constructive way forward, particularly if it embodies domain expertise and community metis. Such a panel could help the police identify potential solutions to the challenges they are grappling with in being data-driven. Initial design exploration with the police indicated that they would be open to such a collaboration, particularly if they could receive technical support and expertise that they often lack at an organizational level. Similarly, the case for strong objectivity through community engagement was also made by organizers, who indicated the importance of listening and including the voices of those "who are the most affected… to take direction from them on how we proceed" [P15].

Here, it is also important to note that simply having community participation is not sufficient. In order to maximize objectivity, differing experiences and perspectives are necessary to create more accurate accounts of police and community activity. As Harding states, "weak objectivity cannot identify paradigms." Knowledge created by a powerful group cannot possibly identify patterns of discrimination and bias in their practices, as experienced by marginalized groups of people. If the police and communities are to truly uncover the patterns of mistrust and inappropriate behavior, data must be shaped in such a way that includes the perspectives and knowledge of all stakeholders, particularly the ones most negatively impacted by police work.

Legitimizing Metis and Alternate Forms of Data

Strong objectivity provides a way to view the problem space around data-driven technologies and practices devaluing metis, expertise, and non-traditional forms of data (unstructured, qualitative information) in the decision-making processes of human-services workers. Because such organizations often serve at-risk individuals, researchers have argued that it is imperative to account for the human, social element of mission-driven organizations, particularly since such organizations invest in people rather than profit [Kong, 2008]. Such work requires skills and expertise that is often invisible and not suitable for being captured digitally. My research illustrates the friction between the different types of expertise required to be data-driven in specific contexts. For instance, while the police are not data experts, data analysts are often not experts in policing. This creates a tension in how data use supports policing on-the-ground and how policing practices support data use. This mismatch between data use and domain literacy raises concerns around faulty judgments and inactionability on-the-ground.

Metis, or experiential knowledge, shapes how police make decisions and interact in unpredictable and difficult situations, and is unlikely to be entirely quantifiable. To deal with the bias towards quantitative data for supporting rationality, designs could legitimize qualitative data in the design of data-driven technologies. Specifically, designs incorporating more qualitative data forms (*e.g.,* police narratives from incident reports, video footage from CCTV and social media, or community members' accounts through written or audio recorded statements) can help police officers reflect on how they produce data and how that data supports decision-making. For example, my empirical work also suggests value in linking aggregate views of quantitative data to finer-granularity, unstructured case notes. More specifically, designs supporting the collection, exploration, and visualization of *both* qualitative and quantitative data could foster a certain level of integration of metis, or at least consideration of, in data-driven tools and practices. Further, for accountability purposes, police actions cannot be captured in their entirety through just police collected and interpreted data. Thus, I recommend the community be directly involved by providing official methods for documenting their perspectives. Currently, some systems already exist that document police-citizen interactions (*e.g.*, www.copwatch.org); however, they are not officially referenced by police during investigations. Thus, there is a need to officially capture citizen perspectives where the power of shaping narratives does not lie just with policing agencies.

I argue that without considering either police metis or communities' lived experiences, data-driven accountability efforts will likely remain insufficient, or worse, provide a false sense of security in terms of what the data does and does not reflect about police actions. Being data-driven, in its current form, will also have an impact on how well data supports effective decision-making. I contend that legitimizing data sources that emerge from outside the police jurisdiction, non-traditional data types, as well as the metis of both police and communities, is an important way to support strong objectivity in the design of data-driven technologies.

Regulating Policies around Data Collection & Use

While I believe the above guidelines can provide a constructive way forward, these solutions fall short of systemically addressing the problem of often ad-hoc data use by police organizations. I argue that another way to maximize objectivity around police accountability is to regulate data creation and use through policy. As of now, data use within MMPD is not standardized across districts. While police departments in America have been reporting Uniform Crime Reporting (UCR)[9] data to the FBI for decades, there are no standards or regulations around data-driven policing for accountability. What is more striking is that UCR does not provide guidelines or frameworks to collect data beyond crime and around police actions. This is problematic because police agencies, which are already often strapped for resources and expertise, are trying to keep up with the data-driven cultures around accountability, but without the adequate support or guidelines to do it well. Data practices *need* standardization through legislation. Unless data collection and use is standardized and regulated at the local, state, and federal levels,

[9] https://www.fbi.gov/services/cjis/ucr

comparability and actionability will continue to be hindered. In order to support the principles of democratic participation and governmental oversight, it is important for us, as community members, to understand how police's impact is experienced throughout the nation to develop a better picture of the experiences on-the-ground. Regulation of data use can then also provide avenues for evaluating the police's use of data and on-the-ground practices. Through streamlining the kinds of data that are collected and how it is collected, we stand to create a more inclusive and authentic space for action around police accountability.

Furthermore, this dissertation contributes to strong objectivity by exploring how accountability and democratic participation can be strengthened in power-disparate situations by providing sociotechnical solutions to mitigate issues around who is believed more and what kinds of data are believed more. First, this dissertation addresses concerns around who is believed more, particularly when looking at organizational power and marginalized communities. This empirical research reemphasizes how such power dynamics lead to marginalized experiences not being legitimized. Because activist groups are already conducting the work to increase capacity to legitimize marginalized accounts, contentious spaces can confront and mitigate the unequal citizenship that often exists in society by creating a framework for stakeholder participation. Similarly, developing policies and regulations around data usage can provide a foundation for governmental organizations to depict the problem spaces more accurately and consistently on aggregate levels. By applying the concept of strong objectivity to data-driven policing, this research contributes different ways of how strong objectivity can mitigate the power imbalances in knowledge creation and legitimization.

Secondly, this dissertation offers design interventions to mitigate the epistemological limitations of data use in terms of what kinds of data are believed more. Because data-driven technologies and practices are biased towards certain types of information, strong objectivity can also be applied to how different types of information sources support human-centered decision making. My dissertation expands on the concept of strong objectivity in technological contexts to help mitigate these epistemological biases around quantitative data in data-driven cultures. I contend that strong objectivity provides us ways to conceptualize and design for qualitative forms of data and metis to strengthen decision support which data-driven technologies are meant to provide.

While these design implications focus on police accountability specifically, this empirical work can be applied as "strong concepts" to other design spaces as intermediate-level of knowledge [Hook & Lowgren, 2012]. Strong concepts refer to design knowledge that is generative and carries a core design idea that cuts across domains. Subsequently, applying these tenets of strong objectivity in data use would be helpful for domains, where accountability is an important and contentious topic as well as organizations that work with a human-centered mission.

7.2 Designing for Information Transparency

While strong objectivity allows us to explore sociotechnical solutions to help mitigate some of the power disparities around data-driven technologies and practices, there is still the question of what realities data can inherently demonstrate, particularly around sociopolitical issues. No data are truly "raw"; the identification and decisions of what data are to be measured and how data are categorized are political acts, motivated implicitly or explicitly by different values [Crawford, 2013; Ribes & Jackson, 2013; Manovich, 2011]. Values and biases are embodied through the design of systems and practices as police produce and use data [LeDantec et. al., 2009; Voida et. al., 2014; Friedman et. al., 2006; Swenson, 2014]. Here, I define data bias as the influence users and technology have on how data are collected, cleaned, analyzed, presented, and used in decision-making. For example, data bias in the police context can include (1) categories and information that are and are not included in data collection tools, (2) narratives and agendas that are supported by data use and those which are not, (3) biased historical data perpetuating inequities in decision-making tools, and (4) inaccurate and incomplete data leading to erroneous insights.

My research emphasizes the importance of understanding how data can and cannot support police work and accountability. As e-science literature also indicates, lack of contextual information around how, where, and by whom data is created leads to issues in data use and credibility [Faniel et. al., 2013; Faniel & Jacobsen, 2010; Zimmerman, 2007; Rolland & Lee, 2013]. Data provenance and data lineage are key concepts that have been explored in multiple research contexts to understand the origins of data, what happens to data, and where and how data moves over time. The goal here is to create visibility and transparency into the data pipeline for tracing back errors in data analytics. My discussion builds on this concept because more visible data lineage can also be a social mechanism for accountability in how data is produced, by whom, and for whom when dealing with a power disparate context. Subsequently, I contend that in order to address accountability issues around how data are produced and narratives framed, we must design for information transparency in the data production lifecycle (*i.e.,* data collection, analysis and reporting). Here, I use Turilli and Floridi's definition of information transparency as not an ethical principle in itself, but as a "condition for enabling or impairing other ethical practices of principles" [Turilli & Floridi, 2009]. Data transparency becomes a pre-existing condition for police accountability by first establishing *how* data are produced and turned into information for decision-making. In other words, as data are used to hold police accountable, information transparency can be used as a principle to hold data use accountable. Unless the data production lifecycle becomes accessible to and legible by community members, accountability efforts stand to remain insufficient and haphazard. Indeed, accountability cannot stem from collaborations where the power to frame narratives through data remains with the police and the practices that produce such data are obscured or black boxed.

The concept of information transparency could be applied more concretely throughout the phases of data use, including data planning, collection, analysis, and presentation. For instance, to demonstrate more transparency about how data use has been planned, police could provide information regarding which groups of people designed the data policies or cite references. Additionally, transparency around data collection could pinpoint who the data collectors are, at what point the data was input into the police systems, and what the data collection forms entail (explicit categories vs. self-reporting fields). Similarly, for data analysis and presentation, information about the selected constraints (time range, types of crimes, status of reported crimes, etc.) could help consumers of data reflect upon how the constraints of the data could influence the spaces for action being shaped within policing.

The negative impacts around data bias were particularly brought to the forefront in conversations with my participants regarding politically contentious issues of police brutality and accountability. My research indicated that most of the participants do not consider that police's data use *can* be unbiased in its current form. Consequently, the mistrust that communities feel towards the police bleeds into the data that are produced by the police.

Similarly, organizations beyond policing are also contending with bias in being data-driven, particularly where data shapes decision-making and spaces for action. Information transparency as a strong concept should be integrated into design practices around data in other contexts. Information transparency will not solve the problem of biases in data but can provide for ways for both data providers and consumers to be more conscientious about the factors that shape the narratives supported by data-driven decision-making.

7.3 Designing for Mistrust

When speaking about police accountability, the concept of trust came up repeatedly with the stakeholders in my study. While some participants expressed unwavering trust towards the police, most of my participants spoke of police accountability in reference to the absence of trust in police organizations. This issue regarding trusting was not just limited to the contentious relationship police have with communities. Rather, mistrust was referenced as part of a wider cultural phenomenon of "fake news" as well as data violations from large social media companies, which has further eroded the public's trust in democratic institutions and businesses in the United

States [Lazer et. al., 2018; Satariano, 2019]. The empirical evidence presented in this dissertation also raises questions about what it means to rescue data from "fake news".

As academics, who are hyper-aware that bias in technology and data is natural and inevitable, it is unsurprising for bias to come to the forefront of conversations around design and politically contentious topics, which is all the more reason to use bias constructively. When bias can be understood as situated in a larger system, rather than an individual shortcoming, I contend that we as a society can become better equipped to mitigate the disparities that emerge through bias. If mythologies of neutrality and trust are associated with technology and data use, then how can bias and mistrust be framed as constructive driving forces for design around political issues? In what follows, I discuss the role of mistrust in developing design guidelines for data use for police accountability.

While a goal behind the adoption of data-driven technology is to induce more trust and accountability between stakeholders, my empirical results indicate that polices' use of data stands to perpetuate further mistrust as data is perceived as an extension of police's power. Further, considering the larger cultural context of "fake news," mistrust could be characterized as another byproduct of data use. Matthew Carey, an anthropologist, contends that while we, as a society, value trust as a necessary condition for cooperation and functioning of societal interactions, mistrust is often reduced to the absence of trust-- simply the negative consequences of not having trust in something or someone [Carey, 2017]. However, he advocates that mistrust is an appropriate value and attitude to adopt in certain contexts. He supports this argument through an ethnographic case study of people living in the Atlas Mountains of Morocco. Here, social relations are based on the assumed unknowability of another person, meaning that mistrust in social interactions is often common and accepted. Carey's work is highly relevant to apply to design problems in HCI around politically contentious topics. In what follows, I outline the reasons why mistrust is an appropriate value to consider in the design of police accountability and data use practices. Then, I lay out concrete design implications for data creation and use.

Police relations with POC communities in the United States has historically been a contentious topic, as systemically racist policies and laws have been enforced through the police on-the-ground [William, 2015]. Law enforcement agencies have been under intense social scrutiny, and social concerns range from racial bias to a lack of police oversight and accountability to lack of training [Delgado & Stefancic, 2015]. It is also important to note that this scrutiny is in part enabled by data, as critics of police use data to demonstrate the problems in policing [Lum, 2017; Eckhouse, 2017]. For example, Human Rights Data Analysis Group (HRDAG), a nonprofit organization, "applies rigorous science to the analysis of human rights violations" in the policing context [HRDAG, 2016]. HRDAG's US policing project assesses and improves upon police violence data accuracy. Their work is, in part, a response to the lack of systematic data about police violence, which hinders police accountability [Lum, 2017]. Similarly, another Midwest activist organization released a data report in the last year on racial profiling in panhandling arrests; the reason why this report was necessary was exactly because this data had to be collected as an additional and external effort in the community, and this data was not provided (and worse, not available) from the police [James & Leininger, 2019]. While the police collect data for accountability, activists and community members' needs for data are hardly fulfilled by the data police share.

As a response to these criticisms, the US government did advocate for using data-driven strategies. For instance, the Obama administration launched the Police Data Initiative (PDI) and Data-Driven Justice Initiative (DDJ) in 2015 and 2016, respectively, to use data to "increase transparency, build community trust, and strengthen accountability" [Davis et. al. 2016]. As of late-2016, over 120 jurisdictions across the country, including our case study site, had committed to both PDI and DDJ. The adoption rate of these initiatives reaches about 30% of the American population. It is important to note that the future and impact of these Obama-era policies remains unclear under the Trump administration, especially considering these initiatives were never mandatory. Again, unless police data (both around crime and police actions) are collected through a consistent, regulated data frameworks, data use stands to fall short of being actionable for adequate police accountability.

Consequently, researchers, organizers, and most of my participants were concerned about how police can use data to further marginalize and monitor communities of color [Eubanks, 2018; Goyanes, 2018]. It is important to note that while data and technology use stands to perpetuate several kinds of inequalities in society, the scope of this dissertation raised concerns specifically around racial and class-based biases. This work lies in a broader context of marginalization through technology use and also contributes to the work done by Black Lives Matter (BLM) and human rights activist organizations and researchers. While such issues around systemic bias are larger than technology use, I contribute to the area of HCI that advocates using data defensively to support collective action, protest systemic inequalities and in some cases, force accountability [Collins, 2012].

Moving forward, keeping the systemic inequalities and biases within police organizations in mind, I contend that designing for trust is insufficient and perhaps, a bit idealistic than is appropriate in this context. The societal mistrust in police and police data is substantiated from the various perspectives presented in this research and thus, designing for trust assumes an implicit support for systems like police, which, for several stakeholders, simply does not exist. In order to combat this mistrust between communities and police, designing data-driven technologies, practices, and policies with biases and inequalities in mind can be a helpful way forward. If data use is to really improve the relationship between police and marginalized communities, then those communities' concerns and experiences must be addressed through data use.

Considering the groups of people whose values, interests, and biases are embedded in data-driven technologies and practices, and which groups are left out, it is important to design and use technology with the underlying assumption that racial and classist biases likely shapes data sets and practices. This means that more concrete and standardized data around police activity is necessary in order to better understand the experiences of both marginalized communities and the police. What police currently produce are narratives and data around crime in different communities. However, what remains undisclosed is an understanding of what the police do, who they treat as suspects versus community members, how many unarmed citizens are shot by the police, and what the racial make-up of these interactions are. Aggregate data from cities across the United States could help us, as a society, better understand and support the space for action through design, practice, and policies. For example, such aggregate data could include:

- Data on the make-up of police organizations (i.e., police officers' race, ages, education level, years of experience, etc.).
- Data on the race of community members who are pulled over, that is not self-reported by the police (i.e., included in and pulled from driver's licenses)
- Data on how many chargeable offenses were discovered in proportion to how many POC were pulled over.
- Data about police's discretion to let a person go with a warning vs. a ticket.
- Data about how many interactions end in police shootings (with armed and unarmed community members).
- Data about the numbers of guns police confiscate nationally (because several of the police participants' concerns around personal safety revolved around the prevalence of guns in the country).

Racism in policing, to whatever extent, was a significant concern for the majority of stakeholders in my study. Here, I use Atyia Martin's definition of racism, which is a "historically rooted system of dehumanizing power structures and behavior based on ideologies that reinforce the superiority of white people and inferiority of people of color, while harming both" [Martin, 2017]. While it is clear that police have power in our society to disrupt and serve different communities, systemically speaking, it could be argued that police officers are hardly equipped to do better (i.e., poor and unstandardized training, vicarious trauma, insufficient mental health support, legislation that criminalizes social issues, etc.) [Bottner, 1997; Bond, 2014]. Not only can using data defensively, with mistrust in mind, be helpful to legitimize the experiences of marginalized communities, but it can also help legitimize the role police can play in society and the social challenges they face.

Moreover, this research also provides avenues for future research directions in terms of how being data-driven can support different kinds of accountability in ways that can empower marginalized communities and unburden police organizations. For instance, abolitionist organizers in my study discussed alternative ways of conflict resolution and intracommunal accountability because they do not believe police to be a legitimate institution. Inversely, while police participants understood the need to be held accountable, a few officers did indicate that they are often called for civil conflicts that they cannot do much for in terms of legal charges. Here, data can be helpful in providing aggregated views of social issues that perpetuate criminal patterns. One of the frustrations our participants expressed was that while they understand how police actions impact communities negatively, directly addressing the comorbid issues associated with crime, namely poverty, mental illness, and drug addiction are beyond typical police jurisdiction and training. While other scholars argue that police play a role in exacerbating these issues [Bitner, 1976; Jacobs & Britt, 1979], the role of data in confronting police criticisms could help expand the issues police focus on and provide police with alternative ways of addressing these issues.

Chapter 8: Conclusion

In this dissertation, I address the following questions around the limitations and manifestations of biases in data-driven technologies and practices:

1. How do the mythologies of data use manifest as organizational values for stakeholders? How does data stand to rationalize and legitimize spaces for action for these stakeholders?
2. What are the challenges stakeholders experience in using data-driven technologies for shaping accountability?
3. What are key value tensions for accountability between groups of people with large power disparities (*i.e.,* police and marginalized communities)?

Through this dissertation, I have demonstrated how human-services organizations adopting data-driven cultures are grappling with the human-centered issues emerging, as the limitations of quantitative data manifest in serving at-risk people equitably. Subsequently, I reorient the mythologies of objectivity, transparency, and trust to outline guidelines for the HCI community at the intersection of design, practice, and policy. I examined the sociotechnical contexts around data-driven cultures in human-services organizations and how it impacts principles like power, knowledge, and accountability. I examined this topic through three case studies in the nonprofit and policing contexts, where it is important to mind the unpredictable nature of human-beings [boyd & Crawford, 2012; Marshall, 2016; Verma & Voida, 2016]. I demonstrate how HCI is already engaging in the problem space of limitations and challenges of being data-driven and how a more explicit engagement with how the values and biases that shape data use can help us mitigate the power disparities and exacerbation of existing inequalities in society.

8.1 Chapter Summaries

In each chapter, I have demonstrated the following:

Chapter 1

In Chapter 1, I outlined the problem space in HCI around data limitations and how HCI can contribute sociotechnical interventions to mitigate the biases and power that manifest in data use. I also present a case for how the mythologies associated with data use can be reoriented through design, practice, and policy changes.

Chapter 2

In Chapter 2, I situate my research in literature around information management in human-services organizations, data-driven policing, data bias and politics, community informatics, and accountability in government and HCI. I have identified the gaps in these areas of literature and indicated how my research fits at the intersection of this previous work.

Chapter 3

In Chapter 3, I outline the qualitative methods used in this three-phase research. I illustrate how the methodologies used in these case studies enabled me to better understand data-driven technologies and practices, related sociotechnical challenges, and possible design interventions for human-services.

Chapter 4

In Chapter 4, I present results from the first case study of the use of data-driven systems in a nonprofit, human services organization. Here, I have characterized four mythologies of data-driven systems that participants experience as shared organizational values and are core to their trajectory towards a "culture of data." I have also discussed the ways in which being actionable is impeded by a disconnect between the aggregate views of data and the desired "drill down" views of data that would allow the users to understand how to act in a data-driven context. These findings contribute initial empirical evidence for the impact of data-driven technology's epistemological biases on organizations and suggest implications for the design of technologies to better support data-driven decision making by legitimizing non-traditional forms of data.

Chapter 5

In Chapter 5, I have presented results from my second qualitative field study about the adoption of data-driven policing strategies in a Midwestern police department in the United States. Here, I have identified three key challenges police face with data-driven adoption efforts: *data-driven frictions*, *precarious and inactionable insights*, and *police metis concerns*. I demonstrate the issues that data-driven initiatives create for policing and the open questions police agents face. These findings contribute an empirical account of how policing agents attend to the strengths and limits of big data's knowledge claims, as well as helped me develop the problem space for how biases manifest in data-driven policing and how data-driven policing can hinder accountability.

Chapter 6

In Chapter 6, I shared results from the final qualitative field study with various stakeholders around police accountability, including police officers, anti-police activists, and various community members. Here, the results demonstrated that stakeholders' concerns around data use for police accountability revolve around how data stands to be an extension of police's power. These key issues emerged around who has the power to create and legitimize data as knowledge, who gets to utilize that data to receive and allocate resources and services, and finally, how police's data use stands to exacerbate the policing and surveillance of marginalized communities. This work has built upon the previous two case studies to contribute design guidelines for HCI around data practice, design, and policy.

Chapter 7

In Chapter 7, I have analyzed the empirical work to demonstrate the design space around data-driven technology and practices. This research contributes to the calls for the HCI community to consider design implications around the limitations of data use and what it means for the human-beings being served by such organizations (see Table 2).

Table 2: Contributions to Gaps in Literature

Gap in Literature	Contribution
Studying the human experience of data-driven technologies that manifest the computational turn in thinking in organizations.	Offer empirical case studies of the adoption and use of data-driven technologies and the related sociotechnical breakdowns in data use for human-services.
Multiple researchers have called for investigation into the epistemological biases and how they play out in practice within organizations [Boyd & Crawford, 2012; Crawford, 2013].	Identify key epistemological biases in data-driven technologies and how those biases manifest in and impact organizational decision-making actionability.
While HCI researchers have considered value sensitive design in the context of social injustice issues, there is a dearth of results around how value tensions are embodied in relation to power disparities [Dombrowski et. al., 2016].	Identify value tensions that emerge between power disparate users of data in contentious spaces.

As HCI scholars identify the limitations and challenges with data use, researchers call for investigating the design space to mitigate the problems that emerge with technology production and use [Verma & Voida, 2016; Verma & Dombrowski, 2018].	Developed possible avenues for the HCI community to deal with the power disparities that are reflected in data use through applying existing concepts of strong objectivity, information transparency, and mistrust in order to support more robust accountability.
HCI researchers have also called for consideration of how we provide design guidelines situated at the intersection of design, practice, and policy, for actionable change [Jackson et. al., 2014].	Offer design orientations at the intersection of policy, practice, and design to address the sociotechnical issues that arise from data production and use, based on the identified challenges of this research.

In order to address these limitations, I reorient the mythologies around data use (objectivity, transparency, and trust) into pragmatic design guidelines for the HCI community to address the inherent biases in data use. To mitigate the mythology of objectivity, I use the concept of strong objectivity to demonstrate how designers and practitioners could mitigate the power disparities in whose voices are represented in the design, practice, and policies around data. To address the issue of data framing narratives, I advocate for designing with principles of information transparency. Finally, I argue that to improve data's credibility further, particularly in contentious and power disparate contexts, mistrust is a more appropriate value to adopt to guide data practices.

8.2 Broader Implications

There is an increased pressure to produce evidence of impact and outcomes for key stakeholders, particularly for human-services organizations [Snibbe, 2006]. This research echoes the call to account for the human, social element of such mission-driven organizations, particularly because of their investment in people rather than profit [Gillon et. al., 2012]. My research indicates that quantitative data and data-driven technologies fall short of their promise of helping organizations support ground truth, accountability, and actionability. While quantitative data enables organizations to rationalize and legitimize decision-making, my participants expressed serious concern around how such data shapes and constrains spaces for action and accountability. For instance, in Morgan's seminal scholarship around the metaphors for organizational cultures, he compares quantitative data to magic in primitive society, enabling clear-cut decisions to be made in otherwise ambiguous and uncertain situations [Morgan, 1997]. Subsequently, such decision-support, as evident through this dissertation, is rife with manifestations of biases and power disparities. While unpacking the biases that are embedded in data-driven systems may be particularly important for the organizations in my case studies, the kinds of epistemological biases highlighted in this research could be of relevance to other similar organizations across sectors of society.

Ultimately, this research also calls for the recognition that there is a human being who underlies the data. This means that the question of how to act and make decisions based on data becomes a fundamentally moral one. Particularly considering socially contentious topics of police brutality and marginalized experiences, HCI researchers have a strong responsibility to address the design challenge of how to legitimize data that is most meaningful for being actionable, where what it means to be actionable hinges on the moral treatment of individuals.

Without consideration of these human-centered issues around data-use, data-driven efforts will likely remain insufficient, or worse, provide a false sense of security and continue perpetuating inequalities. While I believe HCI is well suited to take this challenge on, this research has constantly highlighted the intertwined nature of technology and the policies around it. Because technology alone cannot address a history of unequal citizenship and power disparities, we, as HCI researchers, must advocate for more robust and relevant policies that consider the ethics and limitations of technology use, especially in sensitive and consequential contexts of poverty and criminal justice. Similarly, because I have conceptualized bias here as a systemic symptom than individual shortcomings in this research, human-centered policies around data limitations are necessary to achieve a more equitable use of data.

8.3 Future Work

This research provides a few routes for possible research. First, future research should explore the potentially varied relationships among quantitative and qualitative data in data-driven decision making and the actionable use of data. Understanding how data supports decision-making across a variety of different organizations could provide a more comprehensive design space for legitimizing non-traditional forms of data.

Second, because this research has called to reconsider the design for qualitative data (structured and unstructured), it would be beneficial to further study how design could support collection of non-typical data types across the entire ecology of information systems. This research challenge has implications for the user interface down to the underlying infrastructure, as well as for the interoperability of these systems. There need to be accessible ways of collecting—thus, validating—non-typical data types so that they stand a chance of making it into aggregations of data in the first place, as well as accessible ways of aggregating qualitative data across multiple systems.. The design implications of this research extend beyond the data-driven tools and implicate the entire pipeline of information management tools that constitute the ecology of systems being adopted by data-driven organizations.

Third, as more organizations implement data-driven strategies, the need for data literacy training is increasingly evident. Data literacy is important for workers throughout the organization (as well as citizens), ranging from the administrators, who make decisions, to workers on-the-ground, who collect and implement data in the communities. I recommend further investigation of how data literacy could be provided to demonstrate the politicized nature of data—how data plays a role in confirming and countering biases.

Lastly, there is more research required around governmental policies that can regulate data-driven accountability for contexts like the police. My research made it evident that designing to combat the limitations of data with the police was not viable; the police do not have incentive (or the capacity) to work on changing their data practices. There is a need to further understand how data requirements could be standardized to enable consistent data around police accountability throughout the country.

Appendix

Scenarios

Police Perspective Scenario:

Police officers across the USA are evaluated individually and organizationally based on certain quantifiable activities (*i.e.,* number of arrests, tickets, warnings, cases closed, etc.). However, police officers on-the-ground as well as experts in the field have emphasized that several components of their activities and interactions with communities are not and/or cannot be documented, but stand to have a positive impact. For instance, an officer's relationship with community members often leads to certain kinds of information that emerge from having familiarity and personal connections with community members. Similarly, certain pieces of knowledge cannot be captured, which can be based in experiences, local information, and relationships.

<u>Design prompt:</u> Keeping what you have said so far about police accountability in mind and this scenario, brainstorm some ideas (through technology or social interventions) that could help improve police and community relationships in the near future. How could technological or social space help accountability to improve in the next five years? What are ways of holding the police accountable in ways that are not possible right now? What are other ways of holding each other within our communities accountable?

Community Perspective Scenario:

Accountability in government is meant to support (a) identifying performance issues in order to (b) mobilize the public and ultimately, (c) create change in policies and practices. Subsequently, using data to improve accountability raises questions about the process of collecting, analyzing, and presenting data. Who gets to decide what data is collected?

How is it analyzed? How is it presented? What isn't being collected and reported? Data use within a governmental organization can often be an obscure and imbalanced process. The decision-making around what data is collected, by whom, how data is analyzed and reported to the community remains inaccessible to community stakeholders. The community often does not have opportunities to contribute their perspectives in the discussion about what data to collect and for what purpose.

<u>Design prompt:</u>

Keeping what you've said so far about police accountability in mind and this scenario, brainstorm ideas about addressing the lack of input and access community members have to the process of data-driven policing. How can community members engage in what and how data is collected and used by police?

References

1. "Gartner Says Worldwide Business Intelligence and Analytics Market to Reach $16.9 Billion in 2016." 3 February 2016. Online at: http://www.gartner.com/newsroom/id/3198917
2. Archbold, C. A. (2012). *Policing: A text/reader*. Sage.
3. Ariel, B., Farrar, W. A., & Sutherland, A. (2015). The effect of police body-worn cameras on use of force and citizens' complaints against the police: A randomized controlled trial. *Journal of quantitative criminology*, *31*(3), 509-535.
4. Ariel, B., Sutherland, A., Henstock, D., Young, J., Drover, P., Sykes, J., ... & Henderson, R. (2016). Wearing body cameras increases assaults against officers and does not reduce police use of force: Results from a global multi-site experiment. *European Journal of Criminology*, *13*(6), 744-755.
5. Baumer, E. P., & Silberman, M. (2011, May). When the implication is not to design (technology). In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*(pp. 2271-2274). ACM.
6. Benjamin, L. M. (2008). Account space: How accountability requirements shape nonprofit practice. *Nonprofit and Voluntary Sector Quarterly*, *37*(2), 201-223.
7. Benjamin, L. M., & Campbell, D. C. (2015). Nonprofit performance: Accounting for the agency of clients. *Nonprofit and Voluntary Sector Quarterly*, *44*(5), 988-1006.
8. Berkovich, M. (2009, April). Perspective probe: many parts add up to a whole perspective. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems* (pp. 2945-2954). ACM.

9. Berry, D. M. (2011). The computational turn: Thinking about the digital humanities. *Culture Machine*, *12*.

10. Bittner, E. (1967). Police discretion in emergency apprehension of mentally ill persons. *Social problems*, *14*(3), 278-292.

11. Blom, J., Viswanathan, D., Spasojevic, M., Go, J., Acharya, K., & Ahonius, R. (2010, April). Fear and the city: role of mobile services in harnessing safety and security in urban use contexts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1841-1850). ACM.

12. Boehner, K., Vertesi, J., Sengers, P., & Dourish, P. (2007, April). How HCI interprets the probes. In *Proceedings of the SIGCHI conference on Human factors in computing systems*(pp. 1077-1086). ACM.

13. Bond, M. (2014). How education impacts police performance. Retrieved September 2, 2017 from http://inpublicsafety.com/2014/07/how-education-impacts-police-performance/

14. Bopp, C., Harmon, E., & Voida, A. (2017, May). Disempowered by Data: Nonprofits, Social Enterprises, and the Consequences of Data-Driven Work. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3608-3619). ACM.

15. Bovens, Mark A. P. 1999. The Quest for Responsibility. Cambridge: Cambridge University Press

16. boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, *15*(5), 662-679.

17. Braga, A. A., & Weisburd, D. L. (2015). Police innovation and crime prevention: Lessons learned from police research over the past 20 years.

18. Brayne, S. (2017). Big data surveillance: The case of policing. *American Sociological Review*, *82*(5), 977-1008.

19. Brooks, D. (2013). What Data Can't Do. Retrieved February 18, 2015 from http://www.nytimes.com/2013/ 02/19/opinion/brooks-what-data-cant-do.html

20. Burkholder, L. (1992). Philosophy and the Computer.

21. Capowich, G. E., & Roehl, J. A. (1994). Problem-oriented policing: Actions and effectiveness in San Diego. *The challenge of community policing: Testing the promises*, 127-146.

22. Carey, M. (2017). *Mistrust: An ethnographic theory*. Hau Books.

23. Carroll, J. M., & Rosson, M. B. (2007). Participatory design in community informatics. *Design studies*, *28*(3), 243-261.

24. Cavaye, J. (2004, October). Orientating government to support community capacity. In *Changing the Way Government Works: New Interests–New Arrangements seminar, Department of Victorian Communities, IPAA Victoria, Australia*.

25. Chan, J. B. (2001). The technological game: How information technology is transforming police practice. *Criminal Justice*, *1*(2), 139-159.

26. Chan, J. B. (2001). The technological game: How information technology is transforming police practice. *Criminal Justice*, *1*(2), 139-159.

27. Clement, A., & Van den Besselaar, P. (1993). A retrospective look at PD projects. *Communications of the ACM*, *36*(6), 29-37.

28. Collins, P. H. (2012). Social inequality, power, and politics: Intersectionality and American pragmatism in dialogue. *The Journal of Speculative Philosophy*, *26*(2), 442-457.

29. Community Oriented Policing Services. (2017). Retrieved September 1, 2017 from http://www.cops. usdoj.gov

30. Community Oriented Policing Services. 2017. Retrieved September 1, 2017 from http://www.cops. usdoj.gov

31. Corbin, J., & Strauss, A. (2008). Basics of qualitative research: Techniques and procedures for developing grounded theory. *Thousand Oaks*.

32. Crawford, K. (2013). The hidden biases in big data. *HBR Blog Network*, *1*.

33. Dabney, D. (2010). Observations regarding key operational realities in a Compstat model of policing. *Justice Quarterly*, *27*(1), 28-51.

34. Davis, R., Austin, R., and Patil, DJ. (2016). Growing Numbers of Communities Are Using Data To Improve Policing and Criminal Justice. Retrieved August 20, 2017 https://www.justice.gov/archives/opa/blog/growing-number-communities-are-using-data-improve-policing-and-criminal-justice

35. Day, P. and Klein, R. (1987), Accountabilities: Five Public Services (London: Tavistock).

36. Delgado, R., & Stefancic, J. (2015). Critical perspectives on police, policing, and mass incarceration. *Geo. LJ*, *104*, 1531.

37. Dombrowski, L., Harmon, E., & Fox, S. (2016, June). Social justice-oriented interaction design: Outlining key design strategies and commitments.

In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems* (pp. 656-671). ACM.

38. Domonoske, C. (2018, November 8). A Boatload Of Ballots: Midterm Voter Turnout Hit 50-Year High. *NPR*. Retrieved March 1, 2019, from https://www.npr.org/2018/11/08/665197690/a-boatload-of-ballots-midterm-voter-turnout-hit-50-year-high

39. Eckhouse. L. (2017). Big data may be reinforcing racial bias in the criminal justice system. Retrieved September 10, 2017 from https://www.washingtonpost.com/opinions/big-data-may-be-reinforcing-racial-bias-in-the-criminal-justice-system/2017/02/10/d63de518-ee3a-11e6-9973-c5efb7ccfb0d_story.html?utm_term=.1624e4a93d3a

40. Erete, S. L. (2013). Empowerment through community crime-prevention technologies. *interactions*, *20*(6), 27-31.

41. Erete, S. L., Miller, R., & Lewis, D. A. (2014, February). Differences in technology use to support community crime prevention. In *Proceedings of the companion publication of the 17th ACM conference on Computer supported cooperative work & social computing* (pp. 153-156). ACM.

42. Ericson, R. V., Ericson, R. V., & Haggerty, K. D. (1997). *Policing the risk society*. Clarendon Studies in Criminolo.

43. Eriksén, S. (2002, October). Designing for accountability. In *Proceedings of the second Nordic conference on Human-computer interaction* (pp. 177-186). ACM.

44. Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

45. Eversole, R. (2011). Community agency and community engagement: re-theorising participation in governance. *Journal of Public Policy*, *31*(1), 51-71.

46. Faniel, I. M., & Jacobsen, T. E. (2010). Reusing Scientific Data: How Earthquake Engineering Researchers Assess the Reusability of Colleagues' Data. Computer Supported Cooperative Work (CSCW), 19(3-4), 355–375. http://doi.org/10.1007/s10606-010-9117-8

47. Faniel, I., Kansa, E., Whitcher Kansa, S., Barrera-Gomez, J., & Yakel, E. (2013). The Challenges of Digging Data: A Study of Context in Archaeological Data Reuse. In Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries (pp. 295–304). New York, NY, USA: ACM. http://doi.org/10.1145/2467696.2467712

48. Ferguson, R., Clow, D., Macfadyen, L., Essa, A., Dawson, S., & Alexander, S. (2014, March). Setting learning analytics in context: overcoming the barriers to large-scale adoption. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge* (pp. 251-253). ACM.

49. Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings, 1972-1977*. Pantheon.

50. Friedman, B., Kahn Jr, P. H., Borning, A., & Huldtgren, A. (2013). Value sensitive design and information systems. In *Early engagement and new technologies: Opening up the laboratory* (pp. 55-95). Springer Netherlands.

51. Gangadharan, S. P. (2012). Digital inclusion and data profiling. *First Monday*, *17*(5).

52. Garicano, L., & Heaton, P. (2010). Information technology, organization, and productivity in the public sector: Evidence from police departments. *Journal of Labor Economics*, *28*(1), 167-201.

53. Garimella, K., De Francisci Morales, G., Gionis, A., & Mathioudakis, M. (2018, April). Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web* (pp. 913-922). International World Wide Web Conferences Steering Committee.

54. Gillon, K., Brynjolfsson, E., Mithas, S., Griffin, J., & Gupta, M. (2012). Business analytics: Radical shift or incremental change?

55. Goldstein, J. (1960). Police discretion not to invoke the criminal process: Low-visibility decisions in the administration of justice. *The Yale Law Journal*, *69*(4), 543-594.

56. Goodwin, I. (2005). The Internet, Organisational Change and Community Engagement: The Case of Birmingham City Council 1. *Prometheus*, *23*(4), 367-384.

57. Goyanes, R. (2018). Data for Black Lives Is an Organization Using Analytics as a Tool for Social Change. Retrieved from http://www.vice.com.

58. Hale, T. N. (2008). Transparency, accountability, and global governance. *Global Governance: A Review of Multilateralism and International Organizations*, *14*(1), 73-94.

59. Hannula, M., & Pirttimaki, V. (2003). Business intelligence empirical study on the top 50 Finnish companies. *Journal of American Academy of Business*, *2*(2), 593-599.

60. Harding, S. (1995). “Strong objectivity”: A response to the new objectivity question. *Synthese*, *104*(3), 331-349.

61. Hardy, E. (2010). Data-driven policing: How geographic analysis can reduce social harm. *Geography Public Safety*, *2*(3).

62. Haskins, R., & Baron, J. (2011). Building the connection between policy and evidence. *Using Evidence to Improve Social Policy and Practice*, *25*.

63. HDRAG. Policing. 2016. Retrieved September 5, 2017 https://hrdag.org/policing/

64. Herman, R. (1994). The Jossey-Bass Handbook of Nonprofit Management and Leadership. *San Francisco*.

65. Irani, L. C., & Silberman, M. (2016, May). Stories we tell about labor: Turkopticon and the trouble with design. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 4573-4586). ACM.

66. Jackson, S. J., Gillespie, T., & Payette, S. (2014, February). The policy knot: re-integrating policy, practice and design in cscw studies of social computing. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (pp. 588-602). ACM.

67. Jacobs, D., & Britt, D. (1979). Inequality and police use of deadly force: An empirical assessment of a conflict hypothesis. *Social Problems*, *26*(4), 403-412.

68. James, S. & Leininger, N. (2019). Policing Panhandling: Implications of Race and Housing Status in Indianapolis, 2015-2018. Answer Indiana.

69. Joh, E. E. (2014). Policing by numbers: big data and the Fourth Amendment.

70. Kaufman, D. 2005. Myths and Realities of Governance and Corruption. World Bank Governance Program, Washington, DC.

71. Kidd, A. (1994, April). The marks are on the knowledge worker. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 186-191). ACM.

72. Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 2053951714528481.

73. Kong, E. 2008. The development of strategic management in the non□profit context: Intellectual capital in social service non□profit organizations.*International Journal of Management Reviews*, *10*(3), 281–299.

74. Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Schudson, M. (2018). The science of fake news. *Science*, *359*(6380), 1094-1096.

75. Le Dantec, C. A., Poole, E. S., & Wyche, S. P. (2009, April). Values as lived experience: evolving value sensitive design in support of value discovery. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1141-1150). ACM.

76. Levin, B., & Reitzel, J. D. (2018). Report to the nation: hate crimes rise in US cities and counties in time of division and foreign interference.

77. Lewis, S., & Lewis, D. A. (2011). Digitalizing crime prevention theories: How technology affects victim and offender behavior. *International Journal of Criminology and Sociological Theory*, *4*(2).

78. Lewis, S., & Lewis, D. A. (2012, May). Examining technology that supports community policing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*(pp. 1371-1380). ACM.

79. Lind, E. A., & Tyler, T. R. (1988). *The social psychology of procedural justice*. Springer Science & Business Media.

80. Livingston, D. (1997). Police discretion and the quality of life in public places: Courts, communities, and the new policing. *Columbia Law Review*, 551-672.

81. Lum, K. (2017). Limitations of mitigating judicial bias with machine learning. *Nature Human Behaviour*, *1*(7), s41562-017.

82. Madison, J. (1961). The Federalist Papers: No. 51. *New York: New American Library*, *961*, 326.

83. Mair, M., Greiffenhagen, C., & Sharrock, W. W. (2016). Statistical practice: Putting society on display. *Theory, Culture & Society*, *33*(3), 51-77.

84. Manning, P. K. (2001). Technology's ways: Information technology, crime analysis and the rationalizing of policing. *Criminal Justice*, *1*(1), 83-103.

85. Manning, P. K. (2005). The study of policing. *Police quarterly*, *8*(1), 23-43.

86. Manning, P. K. (2008). *The technology of policing: crime mapping, information technology, and the rationality of crime control*. NYU Press.

87. Manovich, L. (2011). Trending: The promises and the challenges of big social data. *Debates in the digital humanities*, *2*, 460-475.

88. Marshall, M., Kirk, D. S., & Vines, J. (2016, May). Accountable: Exploring the Inadequacies of Transparent Financial Practice in the Non-Profit Sector.

In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 1620-1631). ACM.

89. Mattelmäki, T., & Battarbee, K. (2002, January). Empathy probes. In *PDC* (pp. 266-271).

90. McCabe, A. C., Keast, R. L., & Brown, K. A. (2006). Community engagement: towards community as governance.

91. Meyerson, D. (2004). The tempered radicals: How employees push their companies–little by little–to be more socially responsible. *Stanford Social Innovation Review*, 1-23.

92. Miller, J. K., Friedman, B., Jancke, G., & Gill, B. (2007, November). Value tensions in design: the value sensitive design, development, and appropriation of a corporation's groupware system. In *Proceedings of the 2007 international ACM conference on Supporting group work* (pp. 281-290). ACM.

93. Minhaj, H. (2019). Civil Rights Under Trump (Patriot Act with Hasan Minhaj). Retrieved from http://www.netflix.com

94. Morgan, G., Gregory, F., & Roach, C. (1997). Images of organization.

95. Mulgan, R. (2000). 'Accountability': An Ever□Expanding Concept?. *Public administration*, *78*(3), 555-573.

96. Muller, M. J., & Kuhn, S. (1993). Participatory design. *Communications of the ACM*, *36*(6), 24-28.

97. Muller, M. J., Wharton, C., McIver Jr, W. J., & Laux, L. (1997, March). Toward an HCI research and practice agenda based on human needs and social

responsibility. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems* (pp. 155-161). ACM.

98. Nader, L. (1972). Up the anthropologist: perspectives gained from studying up.

99. Northrop, A., Kraemer, K. L., & King, J. L. (1995). Police use of computers. *Journal of Criminal Justice*, *23*(3), 259-275.

100. Nunn, S. (2001). Police information technology: Assessing the effects of computerization on urban police functions. Public Administration Review, 61(2), 221-234.

101. Nunn, S. (2001). Police technology in cities: changes and challenges. Technology in Society, 23(1), 11-27.

102. Peak, K., & Barthe, E. P. (2009). Community Policing and CompStat: Merged, or Mutually Exclusive?. Police Chief, 76(12), 72.

103. Peters, B. G. (2007). Performance-based accountability. Performance accountability and combating corruption, 15-32.

104. Pine, K. H., & Mazmanian, M. (2014, February). Institutional logics of the EMR and the problem of'perfect'but inaccurate accounts. In Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing (pp. 283-294). ACM.

105. Pine, K., & Mazmanian, M. (2015). Emerging Insights on Building Infrastructure for Data-Driven Transparency and Accountability of Organizations. iConference 2015 Proceedings.

106. Purnell, D. (2017, August 23). What Does Police Abolition Mean? The Boston Review. Retrieved March 10, 2019, from http://bostonreview.net/law-justice/derecka-purnell-what-does-police-abolition-mean

107. Reaves, B. (2013). Local Police Departments, 2013: Personnel, Policies, and Practices. Retrieved August 20, 2017 https://www.bjs.gov/content/pub/pdf/lpd13ppp.pdf

108. Reeves, P. L., & Burt, W. L. (2006). Challenges in data-based decision-making: Voices from principals. Educational Horizons, 85(1), 65-71.

109. Ribes, D., & Jackson, S. J. (2013). Data bite man: The work of sustaining a long-term study. Raw data" is an oxymoron, 147-166.

110. Rinehart Kochel, T. (2011). Constructing hot spots policing: Unexamined consequences for disadvantaged populations and for police legitimacy. Criminal justice policy review, 22(3), 350-374.

111. Rogers, J., Ishimoto, M., Kwako, A., Berryman, A., Diera, C. (2019). School and Society in the Age of Trump. Los Angeles, CA: UCLA's Institute for Democracy, Education, and Access.

112. Rolland, B., & Lee, C. P. (2013). Beyond trust and reliability: reusing data in collaborative cancer epidemiology research. In Proceedings of the 2013 conference on Computer supported cooperative work (pp. 435–444). New York, NY, USA: ACM. http://doi.org/10.1145/2441776.2441826

113. Romzek, B. S., & Dubnick, M. J. (1994). Issues of accountability in flexible personnel systems.

114. Ryan, N., Head, B., Keast, R., & Brown, K. (2006). Engaging indigenous communities: Towards a policy framework for indigenous community justice programmes. Social Policy & Administration, 40(3), 304-321.

115. Rydberg, J., & Terrill, W. (2010). The effect of higher education on police behavior. Police Quarterly, 13(1), 92-120.

116. Sadd, S., & Grinc, R. M. (1994). Innovative neighborhood-oriented policing: Descriptions of programs in eight cities. New York: Vera Institute of Justice.

117. Salamon, L. M., Sokolowski, S. W., & List, R. (2004). Global civil society (Vol. 2). Bloomfield, CT: Kumarian.

118. Satariano, A. (2019, January 17). Facebook Identifies Russia-Linked Misinformation Campaign. The New York Times. Retrieved March 1, 2019, from https://www.nytimes.com/2019/01/17/business/facebook-misinformation-russia.html

119. Schedler, A. (1999). Conceptualizing accountability. The self-restraining state: Power and accountability in new democracies, 13, 17.

120. Schuler, D. (1996). New community networks: Wired for change. ACM Press/Addison-Wesley Publishing Co.

121. Scott, J. C. (1998). Seeing like a state: How certain schemes to improve the human condition have failed. Yale University Press.

122. Simonsen, J., & Robertson, T. (Eds.). (2012). Routledge international handbook of participatory design. Routledge.

123. Sklansky, D. A. (2011). The persistent pull of police professionalism.

124. Snibbe, A. C. (2006). Drowning in data. Stanford Social Innovation Review, 4(3), 39-45.

125. Sparrow, M. K. (1988). Implementing community policing (No. 9). Washington, DC: US Department of Justice, National Institute of Justice.

126. Stanley, J. (2013). Police body-mounted cameras: With right policies in place, a win for all. New York: ACLU.

127. Strauss, A., & Corbin, J. M. (1997). Grounded theory in practice. Sage.

128. Suchman, L. (1993). Do categories have politics? The language/action perspective reconsidered. In Proceedings of the Third European Conference on Computer-Supported Cooperative Work 13–17 September 1993, Milan, Italy ECSCW'93 (pp. 1-14). Springer, Dordrecht.

129. Suchman, L. (2002). Located accountabilities in technology production. Scandinavian journal of information systems, 14(2), 7.

130. Sunshine, J., & Tyler, T. R. (2003). The role of procedural justice and legitimacy in shaping public support for policing. Law & society review, 37(3), 513-548.

131. Swenson, J. (2014, March). Establishing an ethical literacy for learning analytics. In Proceedings of the Fourth International Conference on Learning Analytics And Knowledge (pp. 246-250). ACM.

132. Taylor, M. (2007). Community participation in the real world: opportunities and pitfalls in new governance spaces. Urban studies, 44(2), 297-317.

133. Thomas, Paul. 2004. "Accountability." In Handbook of Public Policy, ed. B. Guy Peters and Jon Pierre. London: Sage

134. Turilli, M., & Floridi, L. (2009). The ethics of information transparency. Ethics and Information Technology, 11(2), 105–112. http://doi.org/10.1007/s10676-009-9187-9

135. Tyler, T. R. (2006). Why people obey the law. Princeton University Press.

136. Verma, N., & Dombrowski, L. (2018). Confronting Social Critcisms: Challenges when Adopting Data-Driven Policing Strategies. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. ACM.

137. Verma, N., & Voida, A. (2016, November). On Being Actionable: Mythologies of Business Intelligence and Disconnects in Drill Downs. In GROUP (pp. 325-334).

138. Vetere, F., Davis, H., Gibbs, M. R., Francis, P., & Howard, S. (2006, April). A magic box for understanding intergenerational play. In CHI'06 Extended Abstracts on Human Factors in Computing Systems (pp. 1475-1480). ACM.

139. Voida, A., Dombrowski, L., Hayes, G. R., & Mazmanian, M. (2014, April). Shared values/conflicting logics: working around e-government systems. In Proceedings of the 32nd annual ACM conference on Human factors in computing systems (pp. 3583-3592). ACM.

140. Voida, A., Harmon, E., & Al-Ani, B. (2012, May). Bridging between organizations and the public: volunteer coordinators' uneasy relationship with social computing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1967-1976). ACM.

141. Walker, S. (2001). Police accountability. Belmont, CA: Wadsworth, 199-214.

142. Weisburd, D., Mastrofski, S. D., McNally, A., Greenspan, R., & Willis, J. J. (2003). Reforming to preserve: Compstat and strategic problem solving in American policing. Criminology & Public Policy, 2(3), 421-456.

143. Williams, K. (2015). Our enemies in blue: Police and power in America. AK Press.

144. Winner, L. (1980). Do artifacts have politics?. Daedalus, 121-136.

145. Winograd, T. (1987). A language/action perspective on the design of cooperative work. Human–Computer Interaction, 3(1), 3-30.

146. Yin, R. K. (2017). Case study research and applications: Design and methods. Sage publications.

147. Zimmerman, A. (2007). Not by metadata alone: the use of diverse forms of knowledge to locate data for reuse. Int. J. Digit. Libr., 7(1), 5–16. http://doi.org/10.1007/s00799-007-0015-8

Curriculum Vitae

Nitya Verma

EDUCATION
Human Computer Interaction, PhD. | May 2020
Advisors: Dr. Davide Bolchini & Dr. Lynn Dombrowski
Indiana University

BBA, Computer Information Systems | May 2012
University of Houston - Downtown

RESEARCH INTERESTS
Human computer interaction; interaction design; user experience; user research; big data; data analytics; data-driven policing

PUBLICATIONS
Papers
Verma, N., & Voida, A. (2016, November). On Being Actionable: Mythologies of Business Intelligence and Disconnects in Drill Downs. In *Proceedings of the 2016 Group Conference*.

Chattopadhyay, D., Verma, N., Duke, J., & Bolchini, D. (2018). Design and Evaluation of Trust–Eliciting Cues in Drug–Drug Interaction Alerts. *Interacting with Computers*, *30*(2), 85-98.

Verma, N., & Dombrowski, L. (2018, April). Confronting Social Criticisms: Challenges when Adopting Data-Driven Policing Strategies. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p. 469). ACM.

Conference Posters
Verma, N. & Voida, A., (2015). Values underlying Business Intelligence. Poster presented at Indiana Women in Technology Conference 2015.

Verma, N. & Dombrowski, L. (2017) Aspirations and Negotiations for Data-Driven Policing. Poster presented at Human-Computer Interaction Consortium, 2017.

Extended Abstract
Verma, N. & Voida, A. (2016, May). Mythologies of Business Intelligence. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2341-2347). ACM. (Received Best Paper Honorable Mention.)

Verma, N. (2016, November). Towards Re-Orienting the Big Data Rhetoric. In *Proceedings of the 19th International Conference on Supporting Group Work* (pp. 505-508). ACM.

WORKSHOPS

Digital Ethnography Summer School, RMIT University, Melbourne, AU | February, 2017
Human Computer Interaction Consortium Workshop, Watsonville, CA | June, 2017
Midwest STS workshop, Bloomington, IN | 2015 & 2017
Untold Stories: Working with Third Sector Organizations, CHI, Montreal, Canada | April, 2018

RESEARCH GROUP EXPERIENCE

Graduate Research Assistant, August 2014 – August 2019
Department of HCI
School of Informatics and Computing
Indiana University, Indianapolis

- Co-advised by Dr. Lynn Dombrowski | March 2016 – August 2019
- Advised by Dr. Davide Bolchini | August 2015 – August 2019
- Advised by Dr. Amy Voida | August 2014 – August 2015

TEACHING EXPERIENCE

Associate Instructor, Meaning and Form in HCI (H561) | Spring 2016
Instructor, Introduction to HCI (I270) | Spring 2017

PROFESSIONAL EXPERIENCE

Facebook, Menlo Park, CA
User Experience Research Intern, May 2018 – August 2018

- Designed and conducted foundational research on internal users' needs and work practices using qualitative methods.
- Managed stakeholder requirements and expectations for research impact.
- Presented and shared research findings through write-ups and presentations.

HCSS, Sugarland, TX
Technical Support Analyst, January 2014 – August 2014

- Provide HCSS- software, including construction bidding, job management, equipment management and more, support via telephone, screen sharing, email and one-on-one training.
- Performed diagnostics and troubleshooting for customers under high pressure situations, i.e. million dollar bids deadlines impending in minutes.

Shipcom Wireless, Inc., Houston, TX
Software Engineer, June 2012 – August 2013

- Develop logistics/healthcare application for Windows handheld RFID and barcode scanners using Shipcom's CATAMARAN Development Studio. Backend system integration includes SQL, SAP, People Soft, etc.
    - Design SQL databases, write queries, stored procedures, triggers, etc.
- Enhance user experience of Shipcom's proprietary software CATAMARAN for healthcare agencies in regards to asset management, personnel management, patient elopement and alerts system.

    - Researched through on-site observation, interviewing and bridging the gap between development and product teams, and rigorous testing.
- Prepare and help conduct user acceptance testing for projects.
- Deploy solutions and applications to client's production servers and help with the go-live process.

Harris County Department of Education, Houston, TX
Developer (Internship), January 2012 – June 2012
- Provided an upgrade to the current database system for the Research and Evaluation Department of HCDE in MS Access and SQL. Created forms, queries and reports with the new requirements provided by the users
- Identified application security issues with an application security consultant to identify and solve issues like SQL injection, cross-site scripting, inadequate authentication and authorization, etc. through testing.

SERVICE
Student Representative, HCI Faculty Search Committee | 2015 – 2016
Co-Organizer, Brown Bag Talks, HCI Department | Fall 2015 – 2016
President, Women in Technology | Fall 2015 – Present
Student Q&A Host for HCI Guest Speakers | Spring 2015, 2016 & 2017

SELECTED HONORS
Elite 50 Award | 2016
- For achievements outside of the classroom among 8,100 graduate and professional students.

Best in School of Informatics and Computing | 2016
Honorable Mention for NSF GRFP | 2016
Best Paper Honorable Mention for CHI LBW | 2016
SIGCHI Travel Grant for Group | 2016
Honorable Mention for CHI Paper | 2018

SKILLS
Data Collection
Requirement Definition, Interviewing, Participant Observation, Focus Groups, Surveys, Usability Testing, Task Analysis, Contextual Inquiry, Eyetracking
Data Analysis
Memoing, Inductive Analysis, Coding and Categorization, Affinity Diagramming, Grounded Theory, Eyetracking analysis, quantitative analysis
Design & Prototyping
Sketching, Brainstorming, Storyboarding, User Scenarios, Personas, Wireframing, Low and Hi Fidelity Prototyping with tools like Axure and Balsamiq.
Development
SQL, HTML, CSS, Visual Studios, Microsoft Office Project Server, Access, Visio, SharePoint, Adobe Creative Suite.