

The conundrum of police officer-involved homicides: Counter-data in Los Angeles County

Morgan Currie, Britt S Paris, Irene Pasquetto and Jennifer Pierre

Big Data & Society
July–December 2016: 1–14
© The Author(s) 2016
DOI: 10.1177/2053951716663566
bds.sagepub.com



Abstract

This paper draws from critical data studies and related fields to investigate police officer-involved homicide data for Los Angeles County. We frame police officer-involved homicide data as a rhetorical tool that can reify certain assumptions about the world and extend regimes of power. We highlight the possibility that this type of sensitive civic data can be investigated and employed within local communities through creative practice. Community involvement with data can create a countervailing force to powerful dominant narratives and supplement activist projects that hold local officials accountable for their actions. Our analysis examines four Los Angeles County police officer-involved homicide data sets. First, we provide accounts of the semantics, granularity, scale and transparency of this local data. Then, we describe a “counter data action,” an event that invited members of the community to identify the limits and challenges present in police officer-involved homicide data and to propose new methods for deriving meaning from these indicators and statistics.

Keywords

Data, police, community, hackathon, Los Angeles, infrastructure

Introduction

Now data seems like a dry and boring word, but without it, we cannot understand our world and make it better. How can we address concerns about use of force, how can we address concerns about officer involved shootings, if we do not have a reliable grasp of the demographics and the circumstances of those incidents? [...] Without complete and accurate data, we are left with ideological thunderbolts, and that helps spark unrest and distrust and does not help us get better. —FBI Director James Comey, February 12, 2015. (Comey, 2015)

Newspaper headlines throughout 2014–2015 reported a spate of killings of unarmed African Americans by police officers around the United States. Incidents that took place in Ferguson, Missouri; Baltimore; New York; and Los Angeles, among others, ignited a frustrated public who took to the streets and launched social media campaigns to challenge the violence and

racial profiling by police officials. Local and national administrations responded by implementing task forces and new policies around officer conduct. Yet even with new corrective measures in place, the disturbing trend of racially targeted violence exposed the difficulty of tallying the extent of these killings, thanks to a significant gap of data on the number of police officer-involved homicides (POIHs) across the US. Scholars of Criminal Justice, Public Health, and Public Policy have all documented these failings since the 1970s, (Loftin et al., 2003; Sherman and Langworthy, 1979) but the public outcry throughout 2014–2015 has exposed the data’s alarming incompleteness to a wider

Los Angeles Department of Information Studies, University of California, CA, USA

Corresponding author:

Britt S Paris, UCLA, 290 Charles E Young Dr N Los Angeles, CA 90095, USA.

Email: parisb@ucla.edu



public (Loftin et al., 2003).^{1–3} As FiveThirtyEight reporter Reuben Fischer-Baum notes, the FBI’s Supplementary Homicide Report (SHR)—the federal database of police shootings generally referenced by news reports—is an inadequate accounting of such statistics because it only publishes police homicides declared ‘justified’ (Fischer-Baum, 2014). Furthermore, law enforcement agencies are not mandated to report these deaths, leading to a reporting rate of less than a third. When states do report, the data could be differently collected across states and local jurisdictions (Barry and Jones, 2014; Bump, 2014).

Such elisions in official data come at a time of so-called data deluge as we increasingly turn to data as a mechanism for solving societal problems. This impulse was on display in one of President Barack Obama’s first Memorandums in office. The Transparency and Open Government initiative in 2009 committed to “unprecedented levels of openness,” most visibly through a website of federal databases, data.gov (Obama, 2009). The website’s thousands of executive agency datasets are available without fees and with minimal licensing restrictions; they provide a window into government processes such as budgeting, environmental oversight, and scientific research. Providing a rich set of resources for research and technological innovation, the website also promises greater insight into government procedures.

The lack of data on POIHs, however, reveals that there are ongoing gaps in the government’s transparency efforts. Despite the enormous apparatus our government invests in other types of data collection, data.gov currently contains no downloadable national database of POIH data and only links to a website maintained by the Department of Justice, where the data is not easily accessible or downloadable.⁴ While Obama’s Police Data Initiative is a recent step towards remedying this situation, official information on such killings remains fragmentary and difficult to find (Smith and Austin, 2015). National-level data are overall interested in measures of accountability, yet the ellipses in these datasets seem primarily to result from the lack of a data assemblage that would support the consistent collection and recording of data, as well as the dissemination of the data that is collected, even though these large organizations would ostensibly have the resources and labor power to oversee efficient data production.

Some of the best-kept statistics on national POIH are not government-based but collected by activist groups and newspapers. Two of the largest, KilledByPolice.net and Fatal Encounters, are civilian efforts. Operation Ghetto Storm, published by the Malcolm X grassroots committee, released a 2012 report using statistical information from local police departments on police killings of African Americans

in the U.S. The Center for Policing Equity at UCLA similarly collects and analyzes information on police-civilian encounters, studying racial profiling as one of four primary areas of concern. Recently, both *The Guardian* and *The Washington Post* have also established their own national counts on POIH in the U.S.⁵

The data on homicides collected by law enforcement and civic groups provides a case to examine the cultural and political dimensions of such statistics and to position this data within power struggles to own, manage, and share it among different groups. This approach is in line with the growing area of research called critical data studies, a body of scholarship that shares concerns with critical informatics, stactivism, critical making, and critical information studies. Critical data studies seeks to explore data as situated in complex assemblages of action, from data collection and categorization to its subsequent cleaning, storing, and dissemination. This framework also considers how this data is then granted meaning and value as it becomes operable in different situations. Using this theoretical understanding of data, this research then asks how we can move beyond dissecting and analyzing POIH data towards understanding it as a lever of political action. We have found through our research that community groups are concerned with augmenting the current gaps in local and national data collecting and publishing this data themselves. Through mixed-methods data collection, analysis, and visualization, activists can provide an alternative to the official statistical ellipses and foster a call to political action.

In what follows, we lay out existing research conducted within critical data studies and related fields and situate our research on POIH data within this literature. We present a critical analysis of existing Los Angeles County POIH data to understand how the events and evidence surrounding a life taken become represented as a metric imbued with rhetorical power. Finally, we describe what we term, after Dalton and Thatcher, a *counter-data action*, a hackathon that invited community groups to examine and re-interpret POIH data. The event outcomes challenge the existing “official” data by examining it and remixing it with POIH data collected by local media and community groups in Los Angeles in an effort to build counter-narratives to the federal accounts.

Situating data in assemblages

A major goal of critical data studies is to understand data as situated in socio-technical systems that surround its production, processing, storing, sharing, analysis, and reuse. Data assemblages, as Kitchin terms these vast systems, are comprised not just of database infrastructures, but also the “technological, political,

social and economic apparatuses that frames their nature, operation and work” (Kitchin et al., 2015). Data assemblages are maintained by the practices that emerge around them, including their production and collation with other data systems, their subsequent distribution by scientific literature or financial markets, and citizen efforts to work with data. A growing body of research examines the labor and political economies entailed in the reproduction of these assemblages, focusing on practices as diverse as meteorological data (Bates, 2015); data produced by for-profit education companies (Williamson, 2016), and order security data (Ajana, 2015). In this literature, the production of data assemblages is not a neutral, technical process, but a normative, political, and ethical one that is contingent and often contested, with consequences for subsequent analysis, interpretation, and action (Kitchin, 2014).

Data assemblages also function as a representation of knowledge, with the ability to shape what we know and what we do based on that knowledge. These systems carry complex epistemological implications that encourage careful consideration of what sorts of knowledge we can derive from data, as well as practical implications—what types of action should be taken based on the knowledge resulting from data. As such, we are encouraged to question how decisions made in the design and use of these tools shape our understanding of the world in which we live.

Literature from critical data studies implicates the following three considerations of data assemblages: an epistemological stance that views data as consisting of material and discursive systems that reify certain assumptions about the world, as well as involve and extend regimes of power; the need to investigate methods and designs of data construction, including choices made about definitions of phenomena, granularity, and scope of the dataset, and a dataset’s position within wider assemblages of support; and finally, the importance of counter-data action, or engagement with a critical framework through practical applications in which individuals actively interrogate data and their relation to it, as well as improve data literacy in communities that have particular stakes in certain data sets.

Data as a rhetorical tool

Critical data studies generally critiques the widely held understanding of data as an objective set of facts that exist prior to ideology, politics, or interpretation (Kitchin, 2014). The common adage “the data speaks for itself” assumes that this data can be used to determine direct meaning that transcends context or domain-specific knowledge.⁶ In contrast, data is here understood as a contingent set of processes that shape the objects represented. Observations are therefore not

simply supported by data, but are generated by them, giving data rhetorical power (Rosenberg, 2013).

The position that data is not a transparent and objective phenomenon but can function as a rhetorical tool, inherently tied to the technological, political, social, and economic infrastructures that sustain it, allows it to be investigated as a real force in the world. Data can, as Ian Hacking puts it, “make people up” once people define themselves and build institutions based on the categories of a data ontology, whether mental health schemas or the census (Hacking, 1986). There is accordingly a double-sidedness to data, as Alain Desrosieres asserts, since it functions both as a description for what we know as well as a basis for action, particularly when the data is inscribed in stable systems and institutions. The more opaque, powerful, and efficient the systems, the more capably the data can resist critique and operate as something real in a reinforcing loop (Rosenberg, 2013). Statistical techniques, scientific proofs, and government records all lend themselves to this type of reality, as the material recordings become a basis for wide agreement and hence objectivity.

Crime statistics in this way have long been a contested source of discursive power. Writing in 1965 Savitz and Johnston called attention to the dangers of rhetoric promoting an “aura of infallibility” that often surrounds crime statistics (1049). The authors discuss the media’s role in not properly conveying the limited nature of crime reporting statistics adequately enough to the public. As this data shapes our understanding of phenomena, we then must ask what it is evidence of and how its rhetorical strength or weakness is constituted by the assemblages in which it is situated.

Methods of infrastructural inversion

Infrastructure studies works in tandem with critical data studies to offer a method for analyzing the decisions made at various steps of data gathering, management, design, and display (Bowker, 2007; Bowker et al., 2010). This literature argues that at any given moment, data are only snapshots that have been composed through the modeling and “cooking” that happens in the algorithmic process of data capture (Bowker, 2007). The production of data is not inevitable; protocols, organizational processes, measurement scales, categories, and standards are designed, negotiated, and debated in the process of data generation. As such, the method of infrastructural inversion proposed by Star and Bowker lays bare the historical development of certain statistical and classificatory tools that then are co-constitutive of how users, and the society in which they are situated, see the world (Star and Bowker, 1999). Infrastructural inversion seeks to

uncover those infrastructures—technical, social, political, and economic systems that facilitate certain types of knowledge—that have become invisible as a result of their efficiency or ubiquity. The relationship of infrastructure and the symbolic realms of power is, for example, seen in Bowker and Star’s 1999 examination of the International Statistical Classification of Diseases and Related Health Problems (ICD) for classifying death (Star and Bowker, 1999). This classification situates disease in the system so that the political, ethical, and social contingencies appear naturalized.

Phil Agre’s potent and portentous 1994 work in *Wired* declared that with the massive increase in data being generated over 20 years ago, methods of data production must become more transparent. However, one will note that this prescription for data transparency has hardly been heeded. Infrastructural inversion must often excavate how the ownership and management of data systems impact the types of treatment or statistical application applied to data sets. Additionally, an analysis must be sensitive to how semantics, or the way that the data is related to definitions, is of utmost importance to the meaning derived from a particular dataset. The ways data are described or labeled, for instance, may invite comparisons between datasets that are not warranted, while more explicit labeling encourages clear and necessary comparisons (Agre, 1996). Transparency of data production, semantics, and the level of information that is captured—its granularity—are all facets that serve as clues about the data’s production. Consequently, an investigation of how data on POIH is generated must examine the assemblages of which it is part, and its relation to technical, economic, and ideological assumptions. These attributes exist in combination with the work that data does in the world and how these can influence different ways of communicating, expressing, and taking action.

The case for counter-data action

History tells us that statistical data can serve as the basis for political activism. In 19th century France and Prussia, social reformers and labor activists worked with civil servants to gather statistics on the conditions of labor to improve workers’ living conditions, unemployment, and hygiene, while in Germany reformers used population statistics to introduce social protections such as disability insurance (Desrosières and Naish, 2002; Hacking, 1987). Yet critical data studies depart from 19th century conceptions of statistics that parsed social and economic issues such as poverty and health as objective, distinct from passion, and polemics. Scholars of critical data scholars, rather, set out to expose “the double role of statistics in

representing as well as criticizing reality” (Desrosières and Naish, 2002). They understand data, per Desrosières, as both a kind of description and basis for action. It is through this lens that data practices can ignite new forms of activism and resistance.

Critical geographers Dalton and Thatcher call acts of resistance to politically dominant datasets *counter-data action* (Dalton and Thatcher, 2014). This notion draws from their work in critical geographic information systems (GIS), an approach that diverges from the conventional view of geographic maps as a model of the world, to an understanding of maps as political and legal claims on reality. Based on this framework, purveyors of Public Participatory GIS engage “counter mapping” as a method of emancipatory action—generally by a community looking to reclaim or denounce external dominance of resources. In a similar vein, purveyors of “statactivism” use this term to describe “emerging forms of collective actions that use numbers, measurements, and indicators as a means of denunciation and criticism” (Bruno et al., 2014). Statactivists might collect and deploy data that does not exist to make a cause more visible—an historic example of this is the case of AIDS activists in the 1980s—or resist or reject official state indicators and benchmarks through original data collection. These practices might analyze different types of data of varying levels of granularity, collected from different organizations which all presume to shed light on the same phenomena, so that the contingent and negotiated aspects of data might come to the fore. These examples are all acts of appropriation and intervention, a means of wresting control of the power of statistics by either decrying certain authoritative metrics or devising new ones.

An example of this type of counter-data action related to crime reporting can be found in Conroy and Scassa’s discussion of a data collection model for sexual assault reporting developed in Philadelphia (Conroy and Scassa, 2015). The authors reflect on the unreliability of sexual assault data and reveal that the data generally only address the issue of bringing about an awareness of the gaps in reporting. The scholars propose a model that involves extensive collaboration between the Philadelphia Police Department and local women’s advocacy groups to attempt a more proper handling of sexual assault reporting. As part of this model, the authors suggested the women’s groups conduct annual reviews of sexual assault reports that were deemed unfounded in order to aid in assessing incidents of mishandling. This model directly addresses the nebulous nature and tenuous understanding of government transparency, especially considering its significant ties to accountability and subsequent abilities to address institutional deficiencies through the promotion of community values (Conroy and Scassa, 2015).

These overarching concepts set an important basis for the community-centered framework we set for our counter-data action.

We draw from this work, along with critical data studies and related literature, the need to engage in counter-data action in which researchers address their own positionality, carefully consider the implications of the data they use, and work to find ways to resist unethical uses of data (Dalton and Thatcher, 2014). We suggest that there are certain affordances in investigating sensitive civic data in relation to the assemblages that sustain and support the data sets, how data is constructed, as well as the effects of the data in society at large as a result of their use or rhetoric surrounding the data. It is also important to ask what questions can and cannot be answered by the data.

As we will discuss in the following sections, we begin by analyzing four POIH data assemblages, then describe an event in which members of the community engaged with these POIH datasets in a unique counter-data action. Our goals at this event were first, to identify the limits and challenges present in the local and federal metrics that are already publicly available on POIH, then to propose new methods for deriving meaning from indicators and statistics. In this way, we seek to lay bare the various apparatuses that negotiate process and deploy POIH.

Setting the stage for our counter-data action are four datasets of POIH data for Los Angeles County. During the hackathon, we explored, remixed, and reinterpreted these four databases. Two are managed at the federal level, the FBI's SHR, the National Center for Health Statistics' (NCHS), National Vital Statistics System (NVSS), and two by local organizations, the Los Angeles Times Homicide Report, and the Youth Justice Coalition (YJC). In the following section, we describe these four datasets and introduce each as a data assemblage with unique elements of institutional, legal, financial, and material support (Kitchin et al., 2015).

POIH databases of Los Angeles County

The first of our datasets is the FBI's Supplementary Homicide Report (SHR). The SHR, the most frequently cited among the federal datasets, was begun in 1962 as part of the more extensive Uniform Crime Reporting (UCR) database that the FBI has maintained for 85 years.⁷ While the older UCR provides annual counts of all recorded homicides in aggregate numbers, the SHR supplements the UCR with *granular* details that provide some context of the event, particularly the victims' relation to offenders (Federal Bureau of Investigation, 2015; Sherman and Langworthy, 1979). These details are manually recorded by local law enforcement agencies on a voluntary form; how the

form is filled out might vary, and data fields are treated as optional (see Figure 1 for the form.). Once completed, the form is then compiled and coded either by the FBI or by state-reporting agencies to produce the statistical data for all counties in the U.S. that report it. By recording information into the form's column labeled "circumstances," the SHR allows agencies to report data on justifiable homicides by law enforcement—coded as "Felon Killed by Police Officer" (code 81). The FBI offers no evidence as to whether it provides additional oversight over the accuracy of the forms.

As is clear from ongoing criticisms, the SHR has very weak institutional, legal, and financial ensembles of support. Because the report is not legally mandated, many states decline to participate. In data released on the SHR in 2003, 18 states have opted out from reporting on this classification during certain years, with Washington D.C., Montana, and Nebraska opting out of reporting at least 12 years and Florida opting out entirely ("Bureau of Justice Statistics UCR and NIBRS Participation," n.d.). Even if a form is submitted, data entry is often incomplete—law enforcement reporting a homicide will not always include demographic data, for instance. According to the Guardian, "In 2011, 31% of SHRs omitted the offender's sex, age and race. When the victim was a black male, basic identifying data on the offender was omitted more often, 39.9% of the time" (McCarthy, 2015b).

SHR's decentralized, bottom-up approach also creates problems with consistency: data gathered from local sources confront variable software and media to make the recordings, differences that are hidden in the aggregate. As a White House press release reported, Camden PD "cobble together 41 systems that have individual value, but are not designed to work together, requiring their beat officers to enter the same data multiple times" (Smith and Austin, 2015). Without standardization, the report cautioned, analysis of these sources may not be meaningful. Additionally, the SHR provides only information based on the initial police investigation, not on subsequent decisions made by prosecutors or courts.

Our second dataset, the NVSS, gathers reports that originate from death certificates by a coroner or medical examiner, as required by law in 36 states ("Easy Access to the FBI's Supplementary Homicide Reports," n.d.; Enten, 2012; Federal Bureau of Investigation, 2015). In contrast to the voluntary SHR, the NVSS is mandatory. To be classified as a POIH, this form must certify manner-of-death as a homicide, then provide additional detail in an open text field that asks the coroner to "describe how the injury occurred". Only if an officer is listed as a perpetrator in this description is the death coded, through the International Classification of Disease-10 codes, as

“Death by legal intervention.” Problems of reliability crop up, however, because the instructions for completing the form do not explicitly indicate that police involvement be mentioned at all, while coroners may not even know if the deceased was involved in an attempted arrest at the time of death. (Loftin et al., 2003) Studies have shown the inadequacy of this data, with underreporting as high as 51% in some cases (Sherman and Langworthy, 1979). The NVSS lack of guidelines for the death certification makes underreporting inevitable. Furthermore, unlike the SHR, the NVSS only provides aggregate data at the county level, obscuring demographic data at the level of each incident. So while NVSS captures the most detail, counting many aspects that the other datasets do not such as measures of victim marital status and educational attainment, it does not make this data public except in aggregate (Quinn, 2014).

The third data set, the *LA Times*’ (LAT) comprehensive Homicide Report, gathers statistics and analysis on all deaths within Los Angeles County. The Report is a part of the LAT Data Desk; it uses, at a very minimum, police reports corroborated with the coroner’s reports, and it sometimes supplements these with investigative reporting on cases when money and time allow. The data for each homicide is displayed publicly online on a dynamic map, as well as in individual posts with description about each death. Each post is organized through statistical data capturing neighborhood in which the death occurred, gender, age, race and ethnicity, cause of death, and whether an officer was involved. The LAT is very interested in questions of access and, as such, their website makes information on these homicides easily accessible (Burghardt, 2014). Their interface combines quantitative numbers on police homicides with accompanying qualitative information found through their investigations. LAT has an FAQ on its website with information about how the Homicide Report data is collected and processed, and each individual post includes the contact information of the author for questions or concern from the public. The LAT’s data is browsable but not downloadable on its website; individuals can request the statistical data, which the LAT provides in the form of an Excel spreadsheet.

SHR, NVSS, and LAT have concerns of much wider scope than the specifics of POIH;⁸ they are exhaustive statistical classification systems that attempt to capture an entire range of phenomena—all homicides or all deaths. Our fourth dataset, a report of the deceased collected by the YJC, in contrast, devotes personnel to explicitly capture POIH data. The YJC is a community organization devoted to issues around incarceration, youth, and race. The organization’s report is a database that uses coroner’s reports corroborated with police reports, and in some cases makes it report based

on interviews with the family of the deceased, as well as eye witnesses and community members in the area where the victim was killed. Accompanying demographic data (age, gender, race), data on the neighborhood and address where the homicide occurred, and date of death, the YJC in some cases also provides a photograph of the deceased and a short description for each incident of police homicide (for example, “Called to mental health facility; officers claimed they shot because Saucedo approached with ‘sharp object.’”) Their website is not as widely known as the LAT Homicide Report, nor do they incorporate any sort of interactive elements to the display of their information, but they do make available for download the PDF that documents this information.

The two county-only datasets by LAT and YJC diverge from the federal datasets in at least four important ways. First, the county data offers more granularity than the federal accounts, with data points capturing the place where the individual was killed and the victim’s name. Second, in contrast to the federal datasets, the local data often uses more than one source to verify and detail the incidents. The local datasets also introduce qualitative information at the level of each death, sometimes as a result of investigatory effort. Finally, the local statistics are dynamic. In controversial cases, both organizations follow and document legal proceedings that take place following the homicide and adjust accounts as new information comes to light, capturing contestations that can occur as a homicide are deemed justified or unjustified. The federal SHR, in contrast, maintains local police agencies as the arbiters of a death’s justifiability. The numbers do not capture any subsequent legal procedures that might prove the contrary (Loftin et al., 2003).

The local data, it should be acknowledged, is limited for research in that it does not scale up: the methods used by these sources are specific to local phenomenon and cannot be analyzed alongside local data from other counties that do not use the same collection methodologies. Yet we found that the local data provide a counter-narrative to the strictly quantitative data found in the federal accounts that naturalizes these recordings as official facts. Local data is a rhetorical tool that shapes how we understand police homicides. As we describe below, the counter-data action entailed considering this smaller-scale local data with the larger federal data troves to find cases in which the data did not match, as well as to bring qualitative, interpretive analysis to bear on our understanding of POIH.

A counter-data action: The POIH hackathon

Our approach lies in pushing the critical understanding of sensitive civic data toward the exploration of new

community-based mixed-method approaches to data analysis. In order to do so, we hosted a civic hackathon during which we engaged with local Los Angeles community groups to re-interpret the data sets and develop meaningful practices around POIH data in accordance with diverse sensibilities, experiences, and needs. We take seriously the claim that “we are our tools,” as the quantification of human features and actions can both enable and place limits on knowledge and action (boyd & Crawford, 2012; Thatcher, 2015). The hackathon, as an example of counter-data action, aimed to allow researchers and citizens to reclaim such tools, participate in their development, and employ them to build alternative discourses and meanings.

We conceptualized the hackathon as a hacker-maker space (HMS), a context that generally focuses on open-access workshops devoted to teaching and learning through creative and technical work (DiSalvo et al., 2012). What distinguishes HMS events from the traditional hackathon is a focus on the material dimensions of the objects of investigation (DiSalvo et al., 2012). During these events, the traditional “creative mis-use” of a piece of technology is realized along with its “hands-on re-construction” (Schrock, 2014). Participants re-construct objects produced from one particular design tradition, but deploy different assumptions and new design and collection methods. Such strategies are proposed, tested, and processed, and, in doing so, simultaneously “learned,” in an act similar to Matt Ratto’s *critical making*. Critical making focuses on interrogations of the socio-technical through tracing back the ways “things are made” in order to produce, re-produce, or imagine various objects (Ratto, 2011).

Among the primary goals of the hackathon was to collectively investigate how the POIH data sets were constructed. During the hackathon, we worked side-by-side with the participants with the goal of understanding the ways different actors collect and organize POIH data—in tabular and visual formats, in coding practices and on social media—in order to imagine how these processes might proceed otherwise. The awareness that we gained through this process of critical making provided us with new ways to independently re-interpret and make sense of the local and federal data on POIH. The rest of this section discusses the different interpretations or sense-making projects developed at the hackathon. We categorize these projects into three main types: data as creative practice; triangulation with different types of qualitative content, primarily social media; and alternate metrics. Our goal was neither to improve the validity and reliability of POIH data, nor to come up with policy recommendations, but to explore the ways in which dissident and creative approaches to data analysis can reveal something new and unexpected about their contingent nature.

Overview of the POIH hackathon

We invited the public to investigate POIH data at the Hackathon on Police Brutality held in February, 2015 at UCLA (see Figure 2). In preparation for the hackathon, we gathered and organized the four data sets on Los Angeles County POIH and made these available in a public Google spreadsheet. We advertised the event widely, attracting nearly 50 individuals, who ranged from students and professors to members of police watchdog groups, for the four-hour event. The event opened with a panel comprising speakers to talk about their experiences in investigating police abuses and brutality. Speakers included lead reporters from the LAT Homicide Report, representatives from the Stop LAPD Spying Coalition, and Andrew Schrock from the University of Southern California’s Innovation Lab. The LAT reporters described in detail how they collect, analyze, and visualize data related to POI homicides in Los Angeles County. Stop LAPD Spying representatives discussed the social and technical challenges that occur when community groups try to make sense of the POIH phenomenon, while Schrock provided some background knowledge about the use of a hackathon event as a form of social action.

After the panel, participants organized themselves into four teams. The Stop LAPD Spying Coalition led a discussion on police officer POIHs in LA County. A second group began analyzing inconsistencies between the federal databases and local databases, while a third group scanned social media networks in order to gain insights about specific discovered unreported cases. The last group worked to visualize the extant POIH data to disseminate to the public.⁹

POIH data as a creative practice

Members of the visualization team devised a multimedia project that combined the addresses of all POIH sites contained in YJC’s data set with Google Street View images of the locations where the events took place. The goal was to display the site-specific images side-by-side in columns and rows (see Figure 3). Before being able to merge the two datasets, we worked closely with YJC members and other hackathons’ participants to create an Excel file with all the locations of the homicides. Indeed, because the YJC dataset originated as a word document manually compiled entry by entry, the counter-data action provided computer science students who helped YJC convert their datasets into a spreadsheet for easier processing.

Once displayed, the POIH spreadsheet gained striking spatial and visual dimensions of locales that are, if not specifically and singularly identifiable, then

sufficiently reminiscent of an everyday milieu. If viewers can identify with these images at all, then it renders more palpable the realities of incidents of POIH occurring at these sites. Furthermore, these images can be analyzed in order to identify common environmental features of POIH sites—whether POIH occur more often in streets with particular architectural features or surrounded by other physical features, in parks, or near certain types of buildings or institutions. The work produced by the visualization team in this way encouraged those at the hackathon to think critically about the spatial dimension of incidents of POIH.

POIH social media data and qualitative content

Some of those in attendance at the hackathon were drawn to the challenge of supplementing the available POIH data with qualitative narratives surrounding the POIH incidents. This group mined social media for any indications of online presence for victims, drawing from the individual names in the YJC dataset. Narrowing the search to 2012, the most recent year of available SHR information, the group searched through social media sites—Facebook, Twitter, Tumblr, YouTube, and personal blogs—to identify online presence and lingering traces of POIH victims.

Three main types of online presence emerged: persistence of the deceased's activity online, in the form of blog posts or Facebook comments they wrote while still alive; sensationalized commentary surrounding the victim, particularly in the form of YouTube comments on videos capturing the incident; and social media memorials dedicated to the deceased, with comments from family members, friends, and others.¹⁰ Though we found at least one of the three types of online social media presence for the majority of the victims searched, there remained a portion with no online trace at all, a phenomenon that cannot be properly explained without further research. This sort of work directly acknowledges and memorializes the deceased, as searching for and curating these details aid in providing a narrative of each victim beyond that of a single numerical entry in a collected database.

POIH data as alternate metrics

As a third outcome, the hackathon focused on understanding the shortcomings of the federal POIH data by comparing various datasets. Accordingly, the data mining team performed an in-depth analysis of the multiple inconsistencies between the LA Times data and SHR data set. For example, the LAT Homicide Report data indicated 38 POIHs in contrast with the 33 reported in SHR for the same year (2012) and in the same geographical area. Initially, the hackathon

participants established the goal of identifying the five unreported cases of 2012 in the SHR. However, the situation revealed a much more complicated nature than first expected. In the SHR, the group found that there were 11 reported homicides that did not match the LAT Homicide Report based on age, gender, and date. In the LAT Homicide Report, there were 18 that could not be accounted for in the SHR based on age, date, and location. Of the documented instances of POIH in both datasets, just 23 POIHs were consistent between the two datasets. The group found that five were very close to matching both databases, but were a year off in age, or reported the death in an adjoining neighborhood. Overall, the information contained in the two datasets was largely inconsistent. Thus, it was observed that discrepancies exist not only in the count of the deceased itself, but also in the details of each account.

Discussion: The social values of data

The focus on engendering critical dialogue and creative reinterpretation of POIH data at the local level is absolutely fundamental to the work we present in this paper. One of the more important outcomes of our exercise in counter-data activism was collectively uncovering some of the “values” leading to the production of the local and federal POIH data sets and how these choices inform any knowledge production based on the data. In this case, the hackathon found that POIH data at the local level produced contrasting indicators of lived reality and then used this finding to help participants understand the choices that shaped the data. The hackathon in this way demonstrated the existence of large incongruities between the datasets and spoke to the contingent choices made by the different organizations that collect, store, and report this data. In this discussion section, we present three examples of differences that came to light over the course of the hackathon—distinctions that affect how knowledge of POIH is produced and ultimately normalized.

First, participants of the alternate metrics group discussed how each dataset reveals the decisions made about what details to collect on each homicide, a choice that involves biases of what should be visible and invisible within a schema. These variances in granularity bounded the analysis at the hackathon by the level of detail and comparability possible across datasets. This constraint, which hindered analysis, became a source of discussion. All schemas capture race, gender, and age, but otherwise categories varied. Only the local data included the incident address and names of deceased. LAT and YJC incident accounts had information on whether victims were intoxicated, whether domestic abuse was involved, and whether

witnesses dispute the account, but such information is not categorized in a schema. The federal datasets have no categories for purportedly non-lethal actions, such as tasing, that may lead to death. Furthermore, all datasets failed to capture certain information, such as statistics about the number of officers who fired, number of bullets fired, or number of bullets hit—all details that could shed light on “on differential ‘kill ratios’” of certain police agencies compared to others (Klinger, 2012).

Second, the social media group found that qualitative details at the incident level of the local data reveal community concerns around particular deaths—concerns that are smoothed over and rendered anodyne when represented solely in statistical form. For example, a Facebook memorial for a young woman includes posts by family and friends dedicated to issues around police violence and advocacy for the mentally ill. One comment describes that the young woman died after “cradling a small ball peen hammer in the lobby of the Asian Pacific Family Center where she was receiving mental health services.” The post goes on, “Witnesses at the scene reported that the young woman was sitting calmly with the hammer in her lap before the police arrived.” The comment then ends with a general call for more awareness during police responses to the mentally ill: “It is also a place for us to advocate on behalf of other people like the young woman so that the ignorance of mental health by law enforcement does not lead to another precious life cut short.” Qualitative details such as this at the incident level can restore the narratives of witnesses and loved ones, making concrete what otherwise becomes a formal abstraction. Highlighting and parsing these details contribute to a continued awareness of the contested nature of the data used to represent these incidents.

Finally, during the general discussion with the panelists at the hackathon, the question arose of whether the datasets were counting the same phenomena. What constitutes a POIH is not simple or clear-cut, and enforcement agencies cannot turn to law to determine what constitutes a POIH. The datasets subtly differ in how they define and circumscribe a complex and often politically fraught phenomenon. The FBI’s SHR, for instance, does not specify the definition of a justified homicide beyond “the killing of a felon by a peace officer in the line of duty” (“Uniform Crime Reporting Handbook,” 2004). Falling out of SHR’s Code 81 classification are “unjustified” homicides, a designation that by definition criminalizes victims of homicides, nor does the definition count deaths that are not directly traceable to homicide by an officer, such as suicide or intoxication, even if these are effects related to the officers’ presence. Nor does Code 81 count deaths on

federal property or by federal agents.¹¹ The NVSS similarly only counts deaths in which the underlying cause was directly attributed to “legal interventions” by an officer.¹² Both the NVSS and SHR assume the victim is a “felon” or “lawbreaker”; such semantics a priori criminalize the deceased, endowing the schema the power to define the dead prior to investigation or trial.

The symbolic power of semantics is also found in the classification of POIH in the ICD, the schema used by the NVSS. The CDC’s ICD-10 uses the same code for homicides by law enforcement as it does acts of war and law enforcement. When there is little distinction between homicides by act of war, a designation widely held to be noble, and those caused by law enforcement, it is possible to assume that the classification is based on a legal, moral, and ethical character of homicides of this nature, showing the contamination of these factors on ICD’s narrative of classification.

Policy makers are looking to correct the failings of POIH data. An executive task force that formed after the shooting of Michael Brown in Ferguson, MO, for instance, recommended that local law enforcement agencies require federal agencies to report when their officers kill someone (McCarthy, 2015a; Obama, 2015). Obama’s Police Data Initiative seeks to remedy the situation with open data standards that clarify and streamline reporting procedures for law enforcement agencies. Through better data and data-driven solutions, FBI Director Comey insists, we can avoid emotionally driven claims and instead deploy “ideological agnostics that look to information to try to solve problems” (Comey, 2015). Yet, our work at the hackathon revealed that any national account of killings by police officers, even one that redresses current failings, makes choices about how to define a POIH, how to capture conflicting accounts, and the granularity of data about the death to include. These remain interpretive matters constituting the socio-technical dimensions of the data in any case.

In sum, data can be employed to encourage a critical conception of events, revealing patterns that suggest alternate paths for accounting for phenomena, particularly to augment governance systems widely characterized as broken, or in serious need of repair. Rather than acting as conclusive evidence, data can be the starting point to ask new questions that come to light, both through reinterpretations and creative practice. Reflections on the particular outcomes of our counter-data action and the overarching themes they relay about the nature of the data in many ways circle back to the major themes of critical data studies discussed throughout the literature review: that data assemblages are fraught with semantic import, that they operate as a set of relations reproduced by interest groups and material infrastructures, and that POIH data can lead

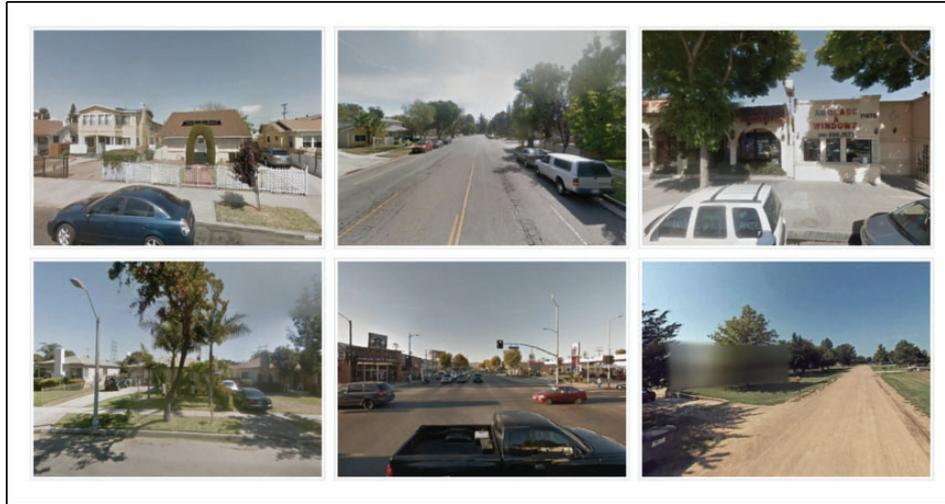


Figure 3. The scraped images of Daniel Schwarz and Visualization Team, from Google Street View, based on address data from the Youth Justice Coalition. Sample of 6 from over 335.

that represent a step in this direction are a branch-off project being conducted by two of the authors currently, as well as an ongoing relationship with the Stop LAPD Spying Coalition. First, two of the authors of this paper are working to build off these findings to inform the creation and implementation of a police harassment reporting mobile application for students on the UCLA campus, a project that will ideally extend the reach and impact of this discussion in a variety of physical and affective ways. Second, the ongoing relationship of the authors with the Stop LAPD Spying Coalition as an outcome of the hackathon has led to a public panel on issues of data, surveillance, and policing, and also a role in the development of a data justice project as part of the coalition's efforts to regain control of data collected about community members by the police. These outcomes are the beginnings of what we hope will continue to develop into beneficial counter-data projects that aid in sustaining a bridge between this type of academic work and the work of community organizers and coalitions.

Conclusion

The dimensions of data explored by critical data studies are key in helping us as a society to understand how data is produced and how it does work in the world, especially in relation to circumstances surrounding the development and use of POIH data. It is difficult to know how data is developed and how, by its very existence—however incomplete—it affects wider understanding of the phenomena these data are supposed to represent or encapsulate. Engaging with this data in counter-data actions at local levels, would, as we have found with our hackathon, remedy widespread problems relating to the rhetoric framing data as

unquestioned truth or a true reflection of the world. These counter-data actions are useful for the communities increasingly affected by data-based governance, as they are a juncture at which citizens can better understand why data is produced and the processes by which it is generated, and that arguments pointing to data are not unassailable.

For example, President Obama's Police Data Initiative seeks to set open data standards that prescribe reporting procedures for law enforcement agencies (McCarthy, 2015a; Obama, 2015). The produced, stored, and disseminated data would be part of a larger assemblage under the auspices of the open data movement—a constellation that involves open formats, open software, open data institutions (such as Code for America and the Open Knowledge Foundation), and open data principles—which will possibly shape the presentation of this data in the future, if not the complete production of the data itself. California is leading in this trend; it published its POIH data online as an open dataset from the California Department of Justice in September 2015. It will be interesting to adherents of critical data studies and data science to follow how this effort continues to develop and its effects on governance and law enforcement.

Yet, even if states follow suit and publish open POIH data, it is likely that data collection practices will continue to differ across states. Many states will still not bother to collect the data because they are not legally bound to. We argue that interrogating the data at the level of infrastructure, production, storage and dissemination and finding discrepancies in these levels in community-based, counter-data actions can frame where there are tensions in the various official state accounts. These differences suggest which

questions should be asked of this data—what are the interests, standards, procedures, and ideologies involved with the construction of the data, and then how can these be meaningfully communicated? These counter-actions can provide an alternate mode of knowledge production in which communities can interact with and interpret qualitative and quantitative data in creative or unexpected ways. As such, we argue that critical data studies expand its gaze to more thoroughly develop modes of counter-data action that take seriously the interpretation and representation of knowledge gained through working with sensitive data sets.

Acknowledgements

The authors thank Leah Lievrouw and Jonathan Furner in the Department of Information Studies at UCLA for supporting the Police Brutality Hackathon. The authors also thank their partners for this event, including Stop LAPD Spying, the Youth Justice Coalition, the Los Angeles Times, Andrew Schrock and Daniel Schwarz as well as the hackathon participants.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: University of California, Los Angeles Department of Information Studies Microsoft Research, FUSE Labs.

Notes

1. Even so, both the press and scholarly research have used federal data about POIH to represent the complex processes that lead to these homicides. Data from the FBI, in particular, is widely cited by the press, though with ample cautions about its validity. USA Today, through an analysis of FBI data, reported on the disconcerting racial demographics of such killings, as “nearly two times a week in the United States, a white police officer killed a black person during a seven-year period ending in 2012.” (Johnson and Hoyer, 2014).
2. A Washington Post analysis based on the FBI data found that officers are increasingly using rifles over handguns in justified fatalities (Bump, 2014).
3. Research in sociology and public policy have likewise relied on these statistics to make assessments and predictions of the use of deadly force overtime and across jurisdictions.
4. The FBI data is available for download through a university website: <http://www.icpsr.umich.edu/icpsrweb/NACJD/archive.jsp>
5. See the websites <http://www.theguardian.com/us-news/ng-interactive/2015/jun/01/the-counted-police-killings-us-database>; <https://www.washingtonpost.com/pb/polices-hootings/>
6. Luciano Floridi claims that the concept of “big data” in particular changes everything about the way we construct and communicate knowledge, because as more data is generated in relation to processes, things and ideas, it changes the information ecology in which humanity exists. Floridi claims that the ubiquity of data grants it a certain type of objectivity that often remains unquestioned—with enough data on a certain process, thing or idea, people take the data as fact and no longer bother to look deeper into how and why the data exists in the first place. As data assumes immanence, it is regarded simply as part of the information ecology in which we exist (Floridi, 2010).
7. These details include jurisdiction; month and year; situation type; the age, sex, and race of the victim and the offender; the weapon used in the homicide; the relationship of the victim to offender; and the circumstance surrounding the offense (e.g. argument, robbery, gang-related).
8. Audits and scholarly reports that cross-reference these two datasets have shown that none contain the same results for the same deaths. Glaringly, according to an audit report by BJS, the FBI’s count was dismally low, capturing 46% of the total from 2003 to 2009, plus 2011; BJS itself captured only slightly more, with 49% during the same time period. The two databases combined missed an average of 263 homicides by law enforcement each year, though not usually the same ones. While BJS ARD data captured more deaths at the aggregate level, it often trailed behind SHR figures once scaled down to individual state reporting. A BJS study from 2007 found California’s SHR figures counted a total 352 justifiable homicides by police officers over three years (i.e. 2003–2005), while the BJS data captured just 160. ARD and SHR counts both missed over half the total between 2003 and 2009. Another study found that nationally, the SHR estimate was 29% larger than the NVSS estimate, but when dialed down to the state-level this trend was reversed, with more deaths reported in the NVSS. (Enten, 2012; Loftin et al., 2003; Mumola, 2007)
9. The results are available at <http://www.poihomicides.org/>
10. The specifics of each of these traces are described in further detail in an earlier conference paper (Currie et al., 2015).
11. This classification leaves out deaths on Native American reservations, in federal prisons, or on military bases; as a count for state law enforcement agencies only, it leaves out actions by federal agencies, such as FBI, DEA, or Marshals Service.
12. Death by legal intervention is defined as “injuries inflicted by the police or other law enforcement agents, including military on duty, in the course of arresting or attempting to arrest lawbreakers, suppressing disturbances, maintaining order, and other legal action.”
13. One example of this kind of work is Public Participation GIS, which engages communities in practices related to GIS data, particularly with the goal of empowering marginalized populations.

References

- Agre P (1996, November) Living data. *Wired* 2(1). Available at: <http://archive.wired.com/wired/archive/2.11/agre.if.html> (accessed 22 July 2016).
- Ajana B (2015) Augmented borders: Big Data and the ethics of immigration control. *Journal of Information, Communication and Ethics in Society* 13(1): 58–78.
- Bowker GC and Jones GC (2014, December 3) Hundreds of police killings are unaccounted in federal stats [News]. Available at: www.wsj.com/articles/hundreds-of-police-killings-are-unaccounted-in-federal-statistics-1417577504 (accessed 22 July 2016).
- Bates J (2015) The secret life of a weather datum. Available at: <https://secretlifeofdata.wordpress.com/> (accessed 2 March 2016).
- Bowker GC (2007) *Memory Practices in the Sciences*. Cambridge, MA: MIT Press.
- Bowker GC, Baker K, Millerand F, et al. (2010) Toward information infrastructure studies: Ways of knowing in a networked environment. In: Hunsinger J, Klastrup L and Allen M (eds) *International Handbook of Internet Research*. Amsterdam, the Netherlands: the Netherlands: Springer, pp. 97–117.
- boyd D and Crawford K (2012) Critical questions for big data. *Information, Communication & Society* 15(5): 662–679.
- Bruno I, Didier E and Vitale T (2014) Statactivism: Forms of action between disclosure and affirmation. *Partecipazione E Conflitto. The Open Journal of Sociopolitical Studies* 7(2): 198–220.
- Bump P (2014, August 15). The number of justifiable police homicides has changed since the 1990s. *The Washington Post*. Available at: www.washingtonpost.com/blogs/the-fix/wp/2014/08/15/how-the-number-of-justified-police-homicide-s-has-changed-since-the-1990s/ (accessed 22 July 2016).
- Bureau of Justice Statistics UCR and NIBRS Participation (n.d.) Bureau of justice statistics UCR and NIBRS participation. Available at: www.bjs.gov/content/nibrstatus.cfm (accessed 22 July 2016).
- Burghardt DB (2014, October 24) Doctoral project on police perpetrated homicides, Los Angeles.
- Comey JB (2015) *Hard Truths: Law Enforcement and Race*. Washington DC: Federal Bureau of Investigation. Available at: www.fbi.gov/news/speeches/hard-truths-law-enforcement-and-race (accessed 22 July 2016).
- Conroy AM and Scassa T (2015) Balancing transparency and accountability with privacy in improving the police handling of sexual assaults. SSRN Scholarly Paper No. ID 2696120.
- Currie M, Paris B, Pasquetto I, et al. (2015) The Police Officer Involved Homicide Database Project. iSchools. iConference Proceedings Newport Beach, CA, 2015. <http://hdl.handle.net/2142/73781>.
- Dalton C and Thatcher J (2014) What does a critical data studies look like, and why do we care? Seven points for a critical approach to “big data.” *Society & Space*. Available at: <http://societyandspace.com/material/commentaries/craig-dalton-and-jim-thatcher-what-does-a-critical-data-studies-look-like-and-why-do-we-care-seven-points-for-a-critical-approach-to-big-data/> (accessed 22 July 2016).
- Desrosières A and Naish C (2002) *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge, MA: Harvard University Press.
- DiSalvo C, Louw M, Holstus D, et al. (2012) Toward a public rhetoric through participatory design: critical engagements and creative expression in the neighborhood networks project. *Design Issues* 28(3): 48–61.
- Easy Access to the FBI’s Supplementary Homicide Reports. (n.d.). National center for juvenile justice. Available at: <http://ojjdp.gov/ojstatbb/ezashr> (accessed 22 July 2016).
- Enten HJ (2012, April 5). Justifiable homicides: the story behind the data and our findings. *The Guardian*. Available at: www.theguardian.com/world/2012/apr/05/justifiable-homicides-guns-data-analysis (accessed 22 July 2016).
- Federal Bureau of Investigation (2015) *Uniform Crime Reports*. Available at: www.fbi.gov/about-us/cjis/ucr (accessed 22 July 2016).
- Fischer-Baum R (2014, August 19). Nobody knows how many Americans the police kill each year. Available at: <http://fivethirtyeight.com/features/how-many-americans-the-police-kill-each-year/> (accessed 22 July 2016).
- Floridi L (2010) *Information: A Very Short Introduction*. Oxford, UK: Oxford University Press.
- Hacking I (1986) Making up people. In: Heller T, Sosna M and Wellberry D (eds) *Reconstructing Individualism: Autonomy, Individuality, and the Self in Western Thought*. Stanford, CA: Stanford University Press, pp. 222–236.
- Hacking I (1987) Prussian numbers 1860–1882. In: Krüger L, Datson LJ and Heidelberger M (eds) *The Probabilistic Revolution, Volume 1*. Cambridge, MA: MIT Press, pp. 377–394.
- Johnson K and Hoyer M (2014, August 15). Local police involved in 400 killings per year. *USA Today*. Available at: www.usatoday.com/story/news/nation/2014/08/14/police-killings-data/14060357/ (accessed 22 July 2016).
- Kitchin R (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. Thousand Oaks, CA: Sage.
- Kitchin R, Lauriault TP and McArdle G (2015) Urban indicators and dashboards: Epistemology, contradictions and power/knowledge. *Regional Studies, Regional Science* 2(1): 43–45.
- Klinger DA (2012) On the problems and promise of research on lethal police violence: A research note. *Homicide Studies* 16(1): 78–96.
- Loftin C, Wiersema B, McDowall D, et al. (2003) Underreporting of justifiable homicides committed by police officers in the United States, 1976–1998. *American Journal of Public Health* 93: 1117–1121.
- McCarthy T (2015a, March 2) Police must report shootings to federal government, suggests Obama taskforce. *The Guardian*. Available at: www.theguardian.com/us-news/2015/mar/02/police-shootings-task-force-21st-century-policing-recommendations (accessed 22 July 2016).
- McCarthy T (2015b, March 21). The counted: inside the search for the real number of police killings in the US. *The Guardian*. Available at: chrome://zotero/content/tab.xul (accessed 22 July 2016).

- Mumola CJ (2007) *Arrest-Related Deaths in the United States, 2003–2005*. Washington, D.C: Bureau of Justice Statistics.
- Obama B (2009). Transparency and open government. memorandum for the heads of executive departments and agencies. *Federal Register* 74(15). Available at: www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment (accessed 22 July 2016).
- Obama B (2015, March) *Remarks by the President after meeting with task force on 21st century policing*. Press Conference, Washington, D.C. Available at: www.whitehouse.gov/the-press-office/2015/03/02/remarks-president-after-meeting-task-force-21st-century-policing (accessed 22 July 2016).
- Quinn B (2014) *The Nation's Two Measures of Homicide*. Washington DC: U.S. Department of Justice, Office of Justice Programs.
- Ratto M (2011) Critical making: Conceptual and material studies in technology and social life. *The Information Society* 27(4): 252–260.
- Rosenberg D (2013) Data before the fact. In: Gitelman L (ed.) *“Raw Data” is an Oxymoron*. Cambridge, MA: MIT Press, pp. 15–40.
- Savitz WM and Johnston N (eds) (1965) *The Sociology of Crime And Delinquency*. New York, NY: Wiley and Sons.
- Schrock AR (2014) “Education in Disguise”: Culture of a hacker and maker space. *InterActions: UCLA Journal of Education and Information Studies* 10(1): 1–25.
- Sherman L and Langworthy R (1979) Measuring homicide by police officers. *Journal of Criminal Law and Criminology* 70: 546–60.
- Smith M and Austin RL (2015, May). Launching the police data initiative [blog]. Available at: www.whitehouse.gov/blog/2015/05/18/launching-police-data-initiative (accessed 22 July 2016).
- Star L and Bowker GC (1999) *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT Press.
- Thatcher J (2015) *Understanding Spatial Media: Locative and Sousveillant Media* (SSRN Scholarly Paper No. ID 2657663). Rochester, NY: Social Science Research Network. Available at: <http://papers.ssrn.com/abstract=2657663> (accessed 22 July 2016).
- Uniform Crime Reporting Handbook (2004) ■■■. Washington, DC: U.S. Department of Justice, Federal Bureau of Investigation.
- Williamson B (2016) Digital methodologies of education governance: Pearson plc and the remediation of methods. *European Educational Research Journal* 15(1): 34–53.

This article is a part of special theme on Critical Data Studies. To see a full list of all articles in this special theme, please click here: <http://bds.sagepub.com/content/critical-data-studies>.