Open Data, the Criminal Justice System, and the Police Data Initiative

10.27.2015

ROBYN CAPLAN, ALEX ROSENBLAT, and DANAH BOYD

Introduction[1]

A number of highly publicized cases showcasing the use and misuse of force by police have sparked new efforts to release public sector information that could shed light on police actions.[2] As news reports about homicides and other violence by police become more prominent, public calls this type of data increase.[3]

And yet, despite a robust data collection program within some areas of the criminal justice system, data on use of force by police remains extremely limited. There is currently no shared inventory of homicides or other uses of violence by police in the line of duty, and no federal law in place to require the collection of such data.[4]

In May 2015, amidst street protests and increased calls for transparency, President Obama launched the Police Data Initiative (PDI), a pilot program which brings together federal government agencies, local police departments, community organizers, and private industry, to encourage the use of data and technology by police departments to build community trust.[5] This initiative, which is part of a broader set of reforms recommended by the President's Task Force on 21st Century Policing, identifies "open data" as a way to facilitate better relations between police and their communities, and calls on departments "to increase transparency" more generally, and "support innovation." As of summer 2015, 26 police departments across the country have signed onto the initiative, and have committed to release over 100 datasets previously unavailable to the public, with each department providing at least 3 datasets on police-citizen interactions. In addition, some jurisdictions, like the state of California, have begun their own initiatives outside of the PDI to make criminal justice data more available and accessible to a wider array of individuals.

This primer addresses a number of questions about the role open data can play within the criminal justice system. What data is being released, and how will it serve private and public

interests? Will sharing existing data do enough to build community trust and transparency, or do new types of data need to be collected? Does data collection and release need to be coordinated across police departments, and, if so, how? What role will data intermediaries – such as journalists and private businesses – play in communicating open data to the public? And how is open data used for advocacy and reform?

What is Open Data? What is Open Criminal Justice Data?

Open data is digital information that is made freely available to the public in a manner amenable to reuse, modification, and sharing for any purpose.[6] Whether data is 'open' may sound like a yes or no question, but openness is a spectrum – e.g., data may be shared in formats that are easier or harder to analyze.

Some stakeholders argue that data must be "machine-readable," "structured for computability," and/or downloadable in bulk to qualify as truly "open."[7] Most open data is thus released in standard, structured non-proprietary formats (e.g., .CSV, .XML, or JSON) through open data "portals," though some police departments are releasing data directly on department sites.[8]

The integration of open data within U.S. government information policies began on President Obama's first day of office on January 20, 2009, when he expressed his commitment to "openness in government" in part through the timely release of downloadable "high-value data sets" by each federal agency.[9] Between 2009-2013, the U.S. government issued three other policies on open government data, and repeatedly brought open data forward as a tool that can help mobilize public participation, increase transparency, and help drive economic development.[10] However, the relationships between open data, "open government," and transparency are not as clear-cut as the terminology might suggest. Some data opened by government, such as health inspection, weather or transportation data, may not promote accountability in governance, but could serve other functions.[11] For instance, private companies could use such data for the delivery of products and services.[12] Journalists could also use it to inform citizens about local issues not related to government transparency, such as traffic conditions.[13]

The amount of data being released by criminal justice agencies is large and has a wide variety of potential uses, even though certain key data is still unavailable. The Sunlight Foundation's growing "inventory of public and private-produced criminal justice data" reveals a significant amount of data and information (the database currently points to over 8000 documents cumulatively) being released by federal, state, and municipal agencies including administrative, financial and demographic data, crime and arrest statistics, calls for service, court and trial information, probation data and, in limited cases, uses of force.[14]

But should all of this data be considered open? And, does it contribute to the goals of transparency and community trust espoused in recent policies for police reform? The vast majority of this data is not on police-citizen interactions and does not fit current definitions of open data. Though open data users prefer machine-readable formats, like .CSV, most of the data included within the Sunlight Foundation's inventory is being released in PDF and HTML. Some proponents of open data have said that publishing data in PDF actually signifies an unwillingness of agencies to disclose information to the public.[15] Others have argued police departments may lack the technical capacity and know-how to understand how different formats affect analysis and re-use. The significant time and labor required to process such formats into data that can be analyzed or re-used can serve as a barrier to effective openness.[16]

The Police Data Initiative (PDI): Actors and Activities

Recent efforts to open more criminal justice data through the Police Data Initiative (PDI) emerged out of testimony provided by witnesses and panel members of the Task Force on 21st Century Policing, who repeatedly called for making data more available "to improve community trust."[17] Though open data within the criminal justice context can refer to a diverse array of data, including crime statistics and call for service data which is more readily available, the Police Data Initiative emphasizes the release of data that can help enhance accountability and transparency, such as police-citizen interaction data (e.g., use of force, and pedestrian stops) that is less available.[18]

Under the umbrella of PDI, police departments have paired with organizations like Code for America, and private technology companies like Socrata, ESRI, and CI Technologies, to provide technical assistance or software to extract and open data to the public. In addition, the Police Foundation is building a national "public safety open data portal" to further increase accessibility.[19] One partnership – between the Charlotte-Mecklenburg PD and the University of Chicago – has analyzed data on key policing activities, and developed more effective internal accountability. Additionally, the White House Presidential Innovation Fellows, and the U.S. Chief Technology Officer and Chief Data Scientist are working with other stakeholders to develop a set of best practices for police departments, an "Open Data Handbook," and an "Open Data and Policing Handbook" with case studies. They are also working with police departments to support community events around the opening of data, such as the recent Summerware hackathon held by the New Orleans Police Department, done in conjunction with Operation Spark, an organization that teaches at-risk youths how to code.[20]

Socrata, an independent vendor, has partnered with most of the PDI cities to build open data portals, including Rutland VT, Montgomery MD, Atlanta GA, New Orleans LA, Dallas TX, and others. Code for America has several "Brigades" – satellite or affiliated Code for America organizations – already in PDI cities, such as Open Austin,[21] Open Oakland,[22] and Hack for

LA.[23] Some jurisdictions, like Montgomery, New Orleans, and Los Angeles[24] are quite advanced in their open data programs, and have released or have concrete plans to release a wide array of open criminal justice data, such as crime statistics, calls for service, and other incident data. Others, like Knoxville, lag behind, without an open data portal, policy, or implementation program as of yet.

While concerns about police use of force drove the creation of the Police Data Initiative, very little use of force data has been released (let alone in machine-readable formats) or is included within publicly available open data implementation plans. Some exceptions to this include Montgomery, which publishes every incident in which a police officer intentionally fires a gun – whether with lethal consequences or not – in a non-machine readable format[25] on its website within 24 hours of the incident.[26] Other police departments publish data through abridged summaries (Los Angeles),[27] annual reports, incident-level data (Austin, TX) or through tables and graphs.[28] Dallas publishes officer-involved shootings (OIS) data on their new open data portal, and Philadelphia publishes OIS at the incident level as HTML tables (their open data portal has not yet launched). Oakland has made use of force data available in .CSV through its open data portal, and makes data from January 1, 2009 to June 30, 2014 available to the public.[29]

Open Criminal Justice Data: Concerns and Challenges

The current skepticism around the potential of open criminal justice data reflects a lack of available data on issues of public concern. Other critiques emerge from the difficulties in coordinating large-scale data collection and release programs, without federally mandated programs to compel disclosure of particular types of open criminal justice data that may serve to standardize and create oversight. Additionally, wide variations in the ability of citizens to truly access and assess the quality of open data highlights problems with information policies that are of more use to some – those who have the knowledge, skills and computing resources necessary to analyze and compute data – than others. This growing gap between users and non-users of open data – coined by some as a newly emerging "data divide" – highlights the need to support additional programs and policies that support community use of open data for purposes of civic accountability. The lack of a clear path between the use and analysis of open data and subsequent public sector reform may limit the effectiveness of open data programs. Without a trajectory from open data and advocacy to policy change, data may be 'open' in theory, but lack impact in practice.

1. Availability and Accuracy of Criminal Justice Data

For issues related to police violence and death in custody, gaps in data collection are determining the amount of open data that can be released. Criminal justice experts note that, though there is a significant amount of data on crime in general, there is no reliable national data on the number of

individuals shot by police each year.[30] This lack of data highlights a key flaw in the plan to use open data to spur reform. If police departments aren't required to track and report data on violence by police, there is no way to ensure that better data on these issues is made publicly available.

The Police Data Initiative is only the most recent effort to coordinate police departments in the collection and reporting of data on police and community interaction, such as uses of force. There have been several attempts in the past to either mandate or facilitate the voluntary reporting of deaths in custody and justified uses of force resulting in death. These include start-and-stop legislative efforts such as the Death in Custody Reporting Program (which ran from 2000 to 2006, and was renewed in 2014),[31] the FBI's Uniform Crime Reporting program (UCR) which collects limited data on people killed by police in justifiable action (though the program is badly named: it is voluntary, rather than truly "uniform" across departments),[32] and the Violent Crime Control and Law Enforcement Act of 1994,[33]which mandated the collection of data on excessive use of force by law enforcement officers (from which the last public report was released in 2001).[34] In most cases, these programs were unsuccessful because they relied on voluntary reporting, lacked penalties, and were ignored by police departments that risked incurring federal oversight through reporting data showing evidence of a pattern or practice of discriminatory policing.[35] In some cases these programs were abruptly discontinued, not kept up-to-date, or were just too limited in the scope of data they requested. FBI Director James Comey has lamented the lack of available data, saying "It's ridiculous that I can't tell you how many people were shot by the police last week, last month, last year."[36]

The FBI's Supplementary Homicide Report released in 2013 – which includes the FBI's estimates of justifiable homicides by police officers – includes data from only 2,700 out of the 22,000 FBI-recognized agencies.[37] In the past, reporting data under a program like the Death in Custody Reporting Act was a condition for eligibility for funds under law enforcement grants.[38] Under the most recent version of the act, if police departments do not report deaths in custody, a 10 percent funding penalty can be enforced at the state attorney general's discretion, though critics of this method doubt that attorney generals consistently enforce these penalties. As part of their most recent data UCR release, Comey said the UCR does plan to begin collecting data about non-fatal shootings between law enforcement and civilians.

Comey acknowledges they "need more law enforcement agencies to submit their justifiable homicide data" to understand what is happening across the country, however, the FBI has not announced any plans to mandate reporting.

Even when data is available, it is often of poor quality or inaccurate. This has been noted by chief statistician for the Department of Justice, Michael Planty, who has said "the FBI's justifiable homicides and the estimates from (arrest-related deaths) both have significant limitations in terms of coverage and reliability that are primarily due to agency participation and measurement

issues."[39] Even in cases of more frequently collected and assessed data types, such as crime statistics, ensuring accuracy and reliability of data remains a concern. In audits of the LAPD and the New Orleans Police Department, done through federal consent decree,[40] investigations into the collection of criminal statistics revealed a pattern of misclassifications of serious violent crimes; for instance, an investigation by the Los Angeles Times found that the LAPD had misclassified nearly 1,200 violent crimes. Similar issues have been spotted by journalists with crime statistics. A story published by DNAInfo showed the NYPD's CompStat statistics on crime are consistently off by "as many as 229 incidents per month."[41] In addition, Stanford professor Jennifer Eberhardt has expressed her concerns regarding the accuracy of existing databases on complaints against police, arguing that citizen complaints against police should not be used as an indicator since citizens will avoid complaining to limit their interaction with police.[42]

The data collected through these programs has repeatedly been shown by independent tracking groups and news agencies to be inaccurate and incomplete. Using Google alerts and crowdsourcing efforts, such as those undertaken by FatalEncounters.org, The Guardian's The Counted, and Mapping Police Violence, has resulted in more robust databases of violence by police in real-time than official sources. A recent investigation by the Marshall Project found these databases are likely more accurate than current federal programs.[43] However, similar efforts on less publicized issues around use of force, stop and frisk, citations, and other more high volume, daily interactions would likely not be as successful. Without the opening of data by government entities, it is unlikely that independent trackers would be able to monitor more high volume interactions through media tracking.

2. Coordinating Open Criminal Justice Data Across Police Departments

Presently, 26 police departments have currently signed on to the most recent effort, the Police Data Initiative – a tiny fraction of the 18,000 state and local law enforcement agencies operating across the country.[44] Of these police departments, there is significant variation in the type and amount of data being collected and reported to federal agencies, as well as the intervals at which data is being released. This variation can lead to problems with comparative analysis and interoperability – the ability of users to combine a large number of similar datasets (to perform analyses on, for instance, uses of force across the U.S.).[45]

Researchers from Sunlight Foundation encountered this challenge in the process of putting together their inventory of U.S. criminal justice open data. Terminology shifts from dataset to dataset. For instance, the terms "close management," "solitary housing unit," and "shu" all referred to the same concept of solitary confinement, but varied across jurisdictions.[46] To make this data interoperable, the researchers had to rely on natural language processing tools to find and match similar terms. It is unclear how this challenge will be dealt with when establishing definitions for politically fraught terms such as "justified homicide" in cases of use of force.

At its base, this is an analytical concern to ensure consistency and interoperability across datasets, but criminal justice data often presents more complex ethical and political issues related to how particular actions are defined or enforced by different participants in the criminal justice system.[47] Relying on a distributed network of police officers and departments to collect and oversee, as well as a network of courts and federal agencies involved in the definition and adjudication of criminal acts, criminal justice data is often imbued with the same values and biases driving the criminal justice system as a whole. In other words, a dataset can present an issue as neat and clean cut, hiding political, social, and ethical dimensions that may remain.

3. Mitigating the Cost of Open Data Programs: From Collection to Release to New Technologies

Though accurate figures are difficult to come by, evidence suggests that open data programs can be costly to implement.[48] Socrata, the private company currently working with most of the PDI cities to build their open data portals, charges an initial set-up fee, as well as an annual maintenance fee. In Los Angeles County, Socrata charged the city $316,608 for the first-year cost and an annual cost of $287,108.[49] In the much smaller Richmond, VA, it is estimated that Socrata would cost the city about $30,000 a year in maintenance costs and licensing fees.[50] Officials involved with the PDI say that police departments can choose less expensive options, such as posting data directly to their site, but such affordable solutions come with their own challenges. A recent call for grant proposals by the U.S. Department of Justice to support "Justice Information Sharing Training and Technical Assistance" reveals one way these programs may be funded, however, there is no direct plan for funding open data programs for local law enforcement agencies through the Police Data Initiative.[51] It is important that, as a pilot program, the Police Data Initiative should determine the cost of open data programs for local law enforcement agencies. Efforts to fund data initiatives through such measures may have more negative effects on building trust between communities and police than is gained through an increase in transparency.

Budgets need to also take into consideration the amount of personnel (and time) that is required to implement and maintain open data programs in addition to costly software such as open data portals, costs associated with data reformatting, complying with open data legislation, liability costs, and public relations costs that may occur when negative information comes to light.[52] In some cases, a PDI participating city may have its own Chief Data Officer, or Chief Information Officer to head the transition, who may be able to tap into a robust open data infrastructure that may exist at the city level. In others, they may have fewer resources, such as limited or non-existent staff with data analysis capabilities, or non-existent infrastructure, requiring that the local police department lead the way to develop open data technologies and policy.

4. Interpreting Open Data: User Groups and Bridging the "Data Divide"

Data should not be thought of as raw or neutral; it is often the result of human decision-making that can be biased or erred, and it takes on additional meaning through the act of analysis and interpretation.[53] If the goal is to increase transparency and build community trust, policymakers must rely on the use of open data by relevant stakeholders that have the resources and skills to responsibly analyze and interpret available data. These stakeholders are often capable of using open data to advance their own interpretations and aims.[54]

Scholars and practitioners skeptical about open data's great promise often point to the gap between data release and the skills required for use and analysis as being the key barrier facing the open data movement. As use of open data becomes more prevalent, we are seeing the emergence of a 'data divide' between those who have the access to and opportunity to use open data and those who do not.[55] Fostering data analysis skills within the general population – the knowledge of how to 'lookup' or find relevant data within datasets, as well as the ability to critically analyze statistics – will become increasingly necessary as interpretations of data compete within the public sphere.

Due to the resources needed to make sense of open data, an ecosystem of third party intermediaries – such as businesses, journalists, and civic tech groups – is growing. This includes private enterprises combining crime statistics with GIS data, such as the crime mapping software SpotCrime, or CrimeReports. Predictive policing software, such as that provided to police departments by the company PredPol, also relies on open data (particularly crime statistics).[56] Part of the PDI's goal in opening data is to "support innovation,"[57] a concept that has previously been tied in other U.S. open data policies with the supporting of small businesses and economic growth. The concern is, however, that private business will be the primary intermediary between the government and public, which could limit open data's civic potential. Private businesses, nonprofits and news agencies have a capacity that individuals and community groups may not, and can put time, money and resources into processing data before use: prepping data through cleaning, standardizing, and organizing, or linking and aggregating different datasets together.

To contribute to transparency, information policies – of which open data is a part – must address every aspect of information 'access' for the widest amount of users.[58] The Police Data Initiative is trying to accomplish this through the hosting of hackathons to encourage use of open data by the public in community events. Already, one such hackathon was hosted in New Orleans, bringing in community members and at-risk youth. Further work needs to be done, however, to make open data more accessible to marginalized communities, who are frequently the target of the types of police actions that erode community trust, but who are rarely the constituents participating in hackathons or interpreting the data about their own communities.

Conclusion: How will data – government or citizen-produced – lead to advocacy, reform, and trust?

While a wide variety of users are beginning to make use of open criminal justice data, the evidence linking open data to transparency and reform is not clear-cut. Indirectly, open data has an impact through its use in exposing government agencies to increased scrutiny by the public.[59] However, there is concern that open data could contribute to information overload, limiting democratic participation as opposed to increasing it.[60] Open data intermediaries – journalists, civil rights groups, and private businesses – will play an important interpretive role for open data, which may shape areas for advocacy.

In a limited number of cases, making criminal justice data available to the public has led directly to reform. Stop-and-frisk data in New York, made publicly available through a series of FOIA requests made by the NAACP, played an essential role in the 2012 ruling that declared the practice of stop-and-frisk unconstitutional.[61] However, though the policy has been reformed to some degree, the practice still overwhelmingly targets Black and Latino populations.[62] Citizen science groups, like the Morris Project, are also using publicly available data for advocacy, using the same NYC "stop and frisk" data to point out racial bias in policing and advocate for residents in affected areas.[63]

As government takes a more active role in making data available, it will become necessary to consider how other forms of data production – such as that of journalists and independent citizens – will be affected by these releases of 'official data.' Independent tracking of police through crowdsourced data collection efforts, such as Fatal Encounters and Mapping Police Violence, and through news media, such as The Guardian's The Counted project, is currently bridging the gap between available data and data the public actually wants. Other types of citizen oversight are also already occurring – in several instances, police misconduct has come to light through iPhone video taken by bystanders. As these forms of citizen oversight bring more attention to police actions, there have been attempts to restrict its production – currently to limited effect. A bill introduced in the Texas House of Representatives to restrict videotaping of police was dropped earlier this year.

Though citizens are bridging this gap, evidence suggests the public does want access to more official data and information about police actions as a way to stimulate reform. A recent poll in California conducted by the ACLU found 79 percent of respondents in favor of increasing access to investigations into police misconduct (if there has been wrongdoing), with another 64 percent believing the public should have access anytime a police officer has been accused.[64] The solution, however, may not be as simple as just opening data. Improving accountability and transparency in policing means open data programs must also work on increasing accuracy and collection of data, improving data literacy, coordinating open data releases and, most importantly, ensuring there are clear ways to translate open data into lasting reform.

Critical Questions

1. Should police departments be required to report data on issues like use of force, to federal agencies?

2. How can we facilitate the collection and release of data on sensitive issues – such as use of force – across police departments? How can we coordinate these datasets between all 18,000 police departments?

3. What mechanisms will help professionalize how police departments 'open' their data? What models exist currently that we can look to for best practices?

4. How can we ensure that data is measuring what we want it to measure? How do we ensure definitions, that can often be fraught with social and political meaning, remain consistently applied across the police departments, and through the entire process of data collection and analysis?

5. How can we ensure meaningful citizen participation in police data initiatives? What data and tools do citizen and advocacy groups need to be able to analyze data to promote transparency and accountability?

6. What is a realistic system of oversight to ensure high quality data is collected and released?

7. What happens when publicly released data differs from the data collected by independent tracking organizations and journalists? How will discrepancies in data be treated, and then fed back into the system of data collection and release?

8. What are the potential challenges to privacy that the release of open data may encounter? How can we address these privacy concerns?

9. How can we ensure that the communities most affected by policing are involved in the process of making meaning from the data about them?

10. As greater emphasis is placed on getting data from law enforcement groups, how do we make certain that citizen data is also leveraged and not restricted?

11. How can data released at the local level build trust between communities and police, if efforts to collect data are not done on a national scale to enable comparative analysis?

12. Can we create common standards for data through a bottom-up, trial-and-error approach? Or do we need standard practices out of the gate for inputting, sharing, and analyzing data, to ensure comparative analysis across jurisdictions and police departments?

13. What is needed to make open data 'useful?' In what ways are open data programs being served by primarily voluntary efforts, like hackathons? What is the impact of sourcing voluntary work for government initiatives?

1   We are very grateful for the strong contributions and insights made by David Robinson, Harlan Yu, Corrine Yu, Seeta Peña Gangadharan, Clarence Wardell, Denice Ross, Patrick Davison, and Angie Waller in the research and production of this primer.

2   McKesson, DeRay. "Opinion: Washington Needs to Tell the Truth about Police Violence." *The Washington Post*, June 16, 2015. https://www.washingtonpost.com/news/powerpost/wp/2015/06/16/opinion-washington-needs-to-tell-the-truth-about-police-violence/.

3   Grothaus, Michael. "The U.S. Doesn't Track Deaths By Police, So Citizens Are Doing It." *Fast Company*, June 18, 2015. http://www.fastcompany.com/3045724/fatal-encounters-crowdsourcing-deaths-by-police.

4   McKesson, DeRay. "Opinion: Washington Needs to Tell the Truth about Police Violence." *The Washington Post*, June 16, 2015. https://www.washingtonpost.com/news/powerpost/wp/2015/06/16/opinion-washington-needs-to-tell-the-truth-about-police-violence/.

5   Smith, Megan, and Roy L. Austin Jr. "Launching the Police Data Initiative." *Whitehouse.gov*, May 18, 2015. https://www.whitehouse.gov/blog/2015/05/18/launching-police-data-initiative.

6   Open Knowledge Foundation. "What is Open Data?" *OpendataHandbook.org*. Accessed August 18, 2015. http://opendatahandbook.org/guide/en/what-is-open-data/.

7   Open Knowledge Foundation. "What is Open Data?" *Open Knowledge*. Accessed August 18, 2015. https://okfn.org/opendata/.

8   "Data.gov." *Data.gov*. Accessed August 18, 2015. http://www.data.gov/.

9   Obama, Barack. "Transparency and Open Government." *Whitehouse.gov*, January 20, 2009. https://www.whitehouse.gov/thepressoffice/TransparencyandOpenGovernment.

10   Orszag, Peter R. "Open government Directive." *WhiteHouse.gov*, December 8, 2009. https://www.whitehouse.gov/open/documents/open-government-directive.

11   Yu, Harlan, and David G. Robinson. "The New Ambiguity of 'Open Government.'" 59 *UCLA Law Review Discourse* 178 (2012): 178-208.

12   The GovLab "About the Open Data 500." *OpenData500*. Accessed August 18, 2015. http://www.opendata500.com/us/about/.

13   Chiweya, Nigel. Interview by Robyn Caplan. Telephone Interview. New York, NY: August 18, 2015.

14   Sunlight Foundation. "Opening Criminal Justice Data." *Sunlight Foundation*. Accessed August 18, 2015. http://sunlightfoundation.com/criminaljustice/.

15   Merrill, Jeremy B. "Purifying the Sea of PDF Data, Automatically." *Open Blog*. Accessed August 18, 2015. http://open.blogs.nytimes.com/2015/04/03/purifying-the-sea-of-pdf-data-automatically/.

16   Castro, Daniel and Travis Korte. "Open Data in the G8: A Review of Progress on the Open Data Charter." *Center for Data Innovation*, March 2015. http://www2.datainnovation.org/2015-open-data-g8.pdf.

17   Eberhardt, Jennifer. Remarks made during the "Presidential Task Force Hearing on 21st Century Policing." *C-SPAN.org*. Accessed August 18, 2015. http://www.c-span.org/video/?323754-1/presidential-task-force-hearing-21st-century-policing.

18   Smith, Megal, and Roy L. Austin Jr. "Launching the Police Data Initiative." *Whitehouse.gov*. May 18, 2015. https://www.whitehouse.gov/blog/2015/05/18/launching-police-data-initiative.

19   "Public Safety Open Data Portal." *PublicSafetyOpenDataPortal.org*. Accessed on August 18, 2015. http://publicsafetydataportal.org/.

20   Operation Spark. "Summerware." *Operation Spark*. Accessed on August 18, 2015. https://operationspark.org/summerware.

21   "About Us." *Open Oakland*. Accessed August 18, 2015. https://www.openoakland.org/about-2/.

22   "About Us." *Open Oakland*. Accessed August 18, 2015. https://www.openoakland.org/about-2/.

23   "Hack for LA." *Meetup.com*. Accessed August 18, 2015. http://www.meetup.com/hackforla/.

24   Both New Orleans and Los Angeles have been under federal consent decree, which means that they have more data available to the public due to audits.

25   "Gaithersburg Barricade/Officer Involved Firearm Discharge Investigated Montgomery County, MD Police – Press Releases." *MyMCPNews.com*. Accessed August 18, 2015. http://www.mymcpnews.com/2014/12/22/gaithersburg-barricadeofficer-involved-firearm-discharge-investigated/.

26   McCarthy, Tom. "The Counted: Inside the Search for the Real Number of Police Killings in the US." *The Guardian*, March 21, 2015. http://www.theguardian.com/us-news/2015/mar/21/police-killings-us-government-statistics.

27   LAPD. "Categorical Use of Force." *LAPDOnline.com*. Accessed August 18, 2015. http://www.lapdonline.org/categoricaluseofforce.

28   Dallas Police Department. "Message from Chief of Police." *DallasPolice.net*. Accessed August 18, 2015. http://dallaspolice.net/ois/ois.html.

29   "Police Use of Force Reports." *Open Oakland*. Accessed August 18, 2015. http://data.openoakland.org/dataset/police-use-force-reports.

30   Lowery, Wesley. "How many police shootings a year? No one knows." *The Washington Post*, September 8, 2014. https://www.washingtonpost.com/news/post-nation/wp/2014/09/08/how-many-police-shootings-a-year-no-one-knows/.

31   Bureau of Justice Statistics. "Data Collection: Deaths in Custody Reporting Program, DCRP." *Office of Justice Programs*. Accessed August 18, 2015. http://www.bjs.gov/index.cfm?ty=dcdetail&iid=243.

32   "Expanded Homicide Data Table 14." *FBI*. Accessed August 18, 2015. https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2012/crime-in-the-u.s.-2012/offenses-known-to-law-enforcement/expanded-homicide/expandedhomicidedatatable14justifiablehomicidebyweaponlawenforcement2008-2012.xls.

33   Violent Crime Control and Law Enforcement Act of 1994, H.R. 3355, 103rd Cong. http://www.gpo.gov/fdsys/pkg/BILLS-103hr3355enr/pdf/BILLS-103hr3355enr.pdf.

34   International Association of Chiefs of Police. "Police Use of Force in America 2001," (2001). http://www.theiacp.org/Portals/0/pdfs/Publications/2001useofforce.pdf.

35   Childress, Sarah. "How the DOJ Reforms a Police Department Like Ferguson – Criminal Justice." *Frontline*, March 4, 2015. http://www.pbs.org/wgbh/pages/frontline/criminal-justice/how-the-doj-reforms-a-police-department-like-ferguson/.

36   Fossett, Katelyn. "The DIY Effort to Count Who Police Kill." *Politico Magazine*, June 9, 2015. http://www.politico.com/magazine/story/2015/06/diy-effort-police-shootings-118796.

37   Hansen, Mark. "Is Google More Accurate Than the FBI?" *The Marshall Project*, July 2, 2015. https://www.themarshallproject.org/2015/07/02/is-google-more-accurate-than-the-fbi.

38   Fossett, Katelyn. "The DIY Effort to Count Who Police Kill." *Politico Magazine*, June 9, 2015. http://www.politico.com/magazine/story/2015/06/diy-effort-police-shootings-118796.html.

39   Lowery, Wesley. "How many police shootings a year? No one knows." *The Washington Post*, September 8, 2014. https://www.washingtonpost.com/news/post-nation/wp/2014/09/08/how-many-police-shootings-a-year-no-one-knows/.

40   Russell, Gordon. "IG Report: NOPD misclassifying many robberies," *The New Orleans Advocate*, August 29, 2014. http://www.theneworleansadvocate.com/home/10101445-123/ig-report-nopd-misclassifying-many.

41   Weiss, M., and Chiwaya, N., "NYPD Crime Statistics Off by up to 229 Incidents a Month, Data Show." *DNAInfo.com*, July 27, 2015. http://www.dnainfo.com/new-york/20150727/civic-center/nypd-crime-statistics-are-off-by-up-229-incidents-month-data-show.

42   Shaw, Emily. "What do we want? Data about police practice!" *Sunlight Foundation*, January 22, 2015. http://sunlightfoundation.com/blog/2015/01/22/what-do-we-want-data-about-police-practice/.

43   Hansen, Mark. "Is Google More Accurate Than the FBI?" *The Marshall Project*, July 2, 2015. https://www.themarshallproject.org/2015/07/02/is-google-more-accurate-than-the-fbi.

44   McCarthy, Tom. "The Counted: inside the search for the real number of police killings in the U.S.," *The Guardian*, March 21, 2015. http://www.theguardian.com/us-news/2015/mar/21/police-killings-us-government-statistics.

45   Open Knowledge Foundation, "What is Open Data," *Open Data Handbook*. Accessed on July 15, 2015. http://opendatahandbook.org/guide/en/what-is-open-data/.

46   Cloud, Daniel. "Illuminating the jargon of criminal justice data with Elasticsearch," *Sunlight Foundation*, June 2, 2015. https://sunlightfoundation.com/blog/2015/06/02/illuminating-the-jargon-of-criminal-justice-data-with-elasticsearch/.

47   Bowker, Geoffrey and Susan Leigh Starr. (1999). *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: MIT. Press.

48   For example, recent reports in Los Angeles County revealed that Socrata – the private company that is providing the open data portal software for many cities and agencies across the U.S. – is charging LA County $316,608 for the first year, with an annual cost for portal upkeep at $287,108. See also: Merl, Jean. "L.A. County Supervisors OK creation of open-data website," *Los Angeles Times*, January 20, 2015. http://www.latimes.com/local/politics/la-me-county-open-data-20150121-story.html

49   Schindler, John R. "L.A. County Supervisors OK Creation of Open-Data Website." *Los Angeles Times*, January 20, 2015. http://www.latimes.com/local/politics/la-me-county-open-data-20150121-story.html.

50   Evans, K. Burnell. "City Launches Online Data Portal." *Richmond Times-Dispatch*, July 5, 2015. http://www.richmond.com/news/local/city-of-richmond/articlebd1a6031-ce75-5630-af21-f425d7166836.html.

51   "Justice Information Sharing Training and Technical Assistance (JIS TTA) Program FY 2015 Competitive Grant Announcement," U.S. Department of justice, Office of Justice Programs, https://www.bja.gov/Funding/14JISTTAsol.pdf.

52   Williams, Matt. "Price Tag Complicates California's Push for a Chief Data Officer, Open Data Portal." *GovTech.com*, June 4, 2015. http://www.govtech.com/state/Price-Tag-Complicates-Californias-Push-for-a-Chief-Data-Officer-Open-Data-Portal.html. See also: Sahuguet, Arnaud and Sangokoya, David. "A 'Calculus' for Open Data." February 9, 2015. https://medium.com/@sahuguet/p-b-d-c-1218ee894400.

53   Smith, Megan and Austin, Roy L. "Launching the Police Data Initiative," *WhiteHouse.gov*, May 18, 2015. https://www.whitehouse.gov/blog/2015/05/18/launching-police-data-initiative.

54   Davies, Tim. "Open data, democracy and public sector reform: A look at open government data use from data.gov.uk" *PracticalParticipation.uk*, 2010. http://www.opendataimpacts.net/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf.

55   Gurstein, Michael. "Open data: Empowering the empowered or effective data use for everyone?" *First Monday*, 2011, Vol. 16, http://firstmonday.org/article/view/3316/2764.

56   PredPol, "Predictive Policing on Gun Violence Using Open Data," Accessed August 26, 2015. http://www.slideshare.net/PredPol/predictive-policing-on-gun-violence-using-open-data.

57   "Launching the Police Data Initiative." *Whitehouse.gov*. Accessed May 18, 2015. https://www.whitehouse.gov/blog/2015/05/18/launching-police-data-initiative.

58   Jaegar, Paul . T., and Jean Carlo Bertot. (2010). Transparency and technological change: Ensuring equal and sustained public access to government information . *Government Information Quarterly*, Vol. 27, 371-376.

59   Davies, Tim. (2010). "Open data, democracy and public sector reform: A look at open government data use from data.gov.uk" *PracticalParticipation.uk*, http://www.opendataimpacts.net/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf.

60   Davies, Tim. (2010). "Open data, democracy and public sector reform: A look at open government data use from data.gov.uk" *PracticalParticipation.uk*, http://www.opendataimpacts.net/report/wp-content/uploads/2010/08/How-is-open-government-data-being-used-in-practice.pdf.

61   Cassidy, John. "The Statistical Debate Behind the Stop-and-Frisk Verdict." *The New Yorker*, August 13, 2013. http://www.newyorker.com/news/john-cassidy/the-statistical-debate-behind-the-stop-and-frisk-verdict.

62   Fermino, J. "De Blasio Exaggerates NYPD Stop-and-Frisk Reforms." *NY Daily News*, April 9, 2015. http://www.nydailynews.com/news/politics/de-blasio-exaggerates-nypd-stop-and-frisk-reforms-article-1.2180048.

63   "About Us - MORRIS JUSTICE: A Public Science Project." *Morris Justice Project*, Accessed October 7, 2015. http://morrisjustice.org/about-us.

64   Friedersdorf, Conor. "California Voters Want Policing Reforms That Politicians Won't Deliver." *The Atlantic*, August 27, 2015. http://www.theatlantic.com/politics/archive/2015/08/california-police-reform/402511/.