Getting our facts straight:
A pivotal juncture for police IT in Oakland

R. K. Belew
rik@ElectronicArtifacts.com
30 March 2013

Crime is an unfortunate feature of every large city, but Oakland certainly seems to have more than its share. From home robberies and car thefts in the hills, to gang-related shootings in especially bad neighborhoods, to increasingly frequent armed robberies everywhere, there is a broad perception that Oakland must do better. To complicate matters, OccupyOakland played into a fissure between Mayor Jean Quan and the Oakland Police department (OPD) that is still tender. Going even farther back, a period of very mixed experience over the last decade and capped by the Negotiated Settlement Agreement (NSA) now being administered by a federal judge, much of Oakland's population is distrustful of its own police department. As a result, folks in the hills are hiring private security patrols and installing surveillance cameras. In other areas, people just don't go outside as much.

In the next few months, central features of a new approach to policing will begin to be implemented in Oakland. Most news-worthy have been recent visits by Robert Wasserman and William Bratton, top cops who have been part of a sea-change in big city policing in New York, Boston and LA. The main headlines have involved Bratton's "stop&frisk" policing strategies in other cities, and their hiring was met with a seven hour City Council meeting February 10 and packed town hall meetings ever since. But their real impact goes much deeper: decisions concerning $5 million on new information technology (IT) for OPD will be made in the next few months.

As a member of CodeForAmerica's Oakland brigade, OpenOakland.org, I have been focused with other "hackers for good" on how more modern IT within OPD and the sharing of some of this data with Oakland's citizens can make us all safer. While the city currently provides a brief window into some OPD data[1], there are serious issues in even this data resource; cities like Chicago and Denver are providing much richer datasets. The new spending – for additional consultants, software and hardware systems, licenses to provide Oakland's data and share in others' data resources – could provide a qualitative improvement within OPD and for OPD's relationship with Oakland citizens, or it could anchor the city and its police into another decade of obsolete systems and data so inconsistent and incomplete that useful analysis is impossible.

Evidence-based policing

There are few domains of social activity, government included, that have not been touched by a new

1 ftp://crimewatchdata.oaklandnet.com/

emphasis on data-based decision-making. "Evidence-based" is an especially common qualifier in medical decision-making, where clinical decisions are based on success and failure in previous cases. As more and more "big data" becomes available, the need to inform decisions that were previously based only on intuition can be connected to available data[2]. CodeForAmerica and broader OpenGovernment activities are trying to bring these same techniques to city, state and federal government.

Policing is a good place to start. The emphasis by both Wasserman and Bratton on CompStat (COMPuter STATistics, police-speak for empirically-based decision-making) has been a constant feature in all of their presentations. They are convinced that by carefully watching patterns in prior crimes, they can make useful predictions about where they are likely to occur next. Wasserman and Bratton need solid data for their plans; then, so will OPD as it executes, monitors and adapts these plans into the future.

The citizens of Oakland need access to accurate data, too. The mutual distrust that is currently evident between city officials, OPD, and the citizens of Oakland are sometimes fed by fundamentally different factual bases. Consider Mayor Quan's focus on "100 blocks" for especially intensive policing, and USC's analysis of the "same" data[3]. As commercial interests begin to scavage crime data (e.g., Trulia.com for real estate pricing), from wherever they can, the potential for data misinterpretation and misuse can only increase. On the other hand, if all these varied stake-holders could come to trust a common reference data set, a common database could also become a basis for building mutual trust.

Opportunities

"Software engineering" used to mean (I speak as an ex-computer science professor who has programming computers since the 1970s) spending enormous amounts of time and money to develop elaborate specifications for software systems, years before a single line of code was written. Often the needs of the corporation had changed so much by the time software was actually written that the process needed to start all over again, and this wasteful cycle repeated itself. Today, much more "agile" approaches are used, with quick-and-dirty initial prototype solutions generated to see how close they are to user needs, and the prototype progressively refined from there. Many of Oakland's IT systems are legacies of old-school software planning, while the quickly changing data becoming available to OPD (and the rest of the city) demands much more agile development methods.

Few conversations about Oakland's crime don't involve guns, but currently OPD provides no data concerning

2 http://www.nytimes.com/2013/03/24/nyregion/mayor-bloombergs-geek-squad.html
3 http://www.sfgate.com/bayarea/article/Jean-Quan-given-incorrect-data-on-100-blocks-3667594.php

whether a gun or knife was used in a crime, let alone what kind of weapons are used. This allows us all to have our favorite "folk theory" regarding gun use generally, their role in crimes, the importance of magazine size, etc. Should schools have armed officers? Should we be training teachers to shoot? The need for clear-eyed, evidence-based approach to these problems is especially clear. And the timing couldn't be better: In the aftermath of the Newtown shootings and in response to the 30,000 gun deaths each year, President Obama signed an executive order allowing gun violence research and treatment like the public health risk it is. Again, the first step is keeping good records.

OPD officers currently spend between a third and a half of their time each day filling out paperwork. Every fast food joint uses technology to make their order takers as efficient as possible; OPD officers deserve at least this consideration. Modern data collection techniques using tools such as computer tablets, voice-recognition, GPS units, audio “Portable Digital Recording Devices” and even cameras, allow immediate acquisition of many basic data facts. This could then allow sworn officers to apply their best efforts to the important bits, interpreting and categorizing the basic facts into a meaningful context.

A central truth emerging from the analysis of big data is that it becomes even more valuable when multiple data sources are combined. In the OPD case, an excellent example of the utility of combining across data sources is provided by two separate data streams, one coming from ShotSpotter and the from OPD's dispatch call (911) records. ShotSpotter uses revolutionary acoustic technology to listen for the special sonic signature of gunshots. It appears that in certain neighborhoods shots are being fired, but 911 is rarely called. Combining these data streams allows the identification of a fundamental disconnect between these neighborhoods and OPD, and allows OPD to focus its resources on just such problem areas.

A subtle but fundamental feature of crime reporting concerns the "unit of analysis." All police departments must provide data to the FBI, and since the 1930s this has been in terms of the Uniform Crime Report (UCR) which associates a single reported item with every major crime incident. For example, if criminal using a gun comes into a store, holds up several customers, robs the cash register, and shoots one of the customers, a single entry reporting this event is passed on to the FBI. There is complicated, hierarchic priority system that determines which of these crimes is considered the major one; all the other details go unreported. Further, if for example the shot customer later dies, the fact that a murder has now occurred requires modification to UCR reports, perhaps months after the crime itself.

For reasons like this, the FBI has also developed a more modern reporting system technology, the National Incident-Based Reporting System (NIBRS). Almost certainly, this modern reporting system would make more

accurate reporting by OPD possible. For example, UCR assumes all rapes are of female victims; NIBRS allows both male and female victims. UCR does not differentiate between attempted and completed offenses; NIBRS does. And in NIBRS, multiple crimes can be reported as offenses within the same incident. As of June 2012, 32 states have been certified to report NIBRS to the FBI[4]. Switching reporting standards like this is never easy, but a move to NIBRS as part of OPD's broader IT overhaul would let Oakland fully participate in modern policing efforts.

There will certainly be new threats as new data is collected and more of it made publicly accessible. For example, as additional information about the guns being used in crimes becomes publicly available, the relative value of these weapons both locally and in remote jurisdictions could be affected. There will also always be a complicated balances between information collected to support public safety, and privacy restrictions. One simple example concerns crimes which, because they involve children, domestic disturbances, or sex crimes, have had some of their details redacted prior to publication. Currently, separation of these sensitive fields from the rest of the data depends on careful manual attention by OPD staff. These activities are performed by people with many other responsibilities, under time pressure, with personnel changing along the line. It is no wonder that facts that should not get out sometimes do, or that overly broad filters are applied, just to be safe. While the basis for any particular redaction can be argued, the implementation of these data filters needs to be carefully analyzed and developed into database queries that are clearly articulated and then securely and reliably applied.

Who to contact

What can we do? I encourage you to contact as many of the decision-makers involved as Oakland and OPD get ready to spend a great deal of money on new information technology systems.

- Central to these decisions will be Oakland's city administrator, Deanna Santana. New CIO/CTO administrators and other key staff have been or soon will be appointed. Their relationship with and support of OPD needs will be critical.

- Mayor Quan and the city Council are our primary conduits for citizen input. Council President Patricia Kernighan and member Libby Schaaf have been particularly engaged in community efforts on just these issues.

- The fact that OPD is under the NSE makes Judge Thelton Henderson and Robert Warshaw, his court-appointed monitor, another important user of the same data streams. The importance of OPD data

4 http://www.jrsa.org/

becomes clear in Warshaw's last report[5]:

- *There will be circumstances in which we will be unable to determine fully the compliance status of a particular requirement due to a lack of data, incomplete data, or other reasons that do not support the completion of our work in a manner consistent with timely reporting.*
- *During the current reporting period, OPD did not produce any summary of data collection or analysis of data, noting a continuing issue with data collection (forms), specifically regarding the selections options for the reason for the stop. This problem significantly affects the value of the data...*
- *... persistent problems in accurately recording the number of arrests made by individual officers. Although the specific problems were identified, the Department "resolved" this issue through a process of entering data by hand. As noted in our previous reports, this temporary fix is significant, but it does not stabilize the system to assure ongoing quality in data collection and storage.*
- *The work on a new database provides another opportunity for the Department to examine these issues.*

- Finally, Oakland has hired two of the nation's top experts to provide expertise on just these issues, Robert Wasserman and William Bratton. Bratton in particular was the pioneer of data-driven policing in New York City, where it had an order of magnitude effect on crime. They will both have very useful input into what systems are purchased, and have staked their reputations on improving Oakland safety. As with every consultant, we need to make sure our questions get answered before they leave town.

But it is Oakland's citizens who are paying for these high-priced consultants, and who will pay for whatever new IT systems they specify. It is a very good time for us to pay particularly close attention.

5 http://www2.oaklandnet.com/Government/o/OPD/o/BureauofInvestigation/OAK039538