The New York City Police Department’s Domain Awareness System

E. S. Levine, a Jessica Tisch, a Anthony Tasso, a Michael Joy a

a New York City Police Department, New York, New York 10038
Contact: evan.levine@nypd.org (ESL), jessica.tisch@nypd.org (JT), anthony.tasso@nypd.org (AT), michael.joy@nypd.org (MJ)

Published Online in Articles in Advance: January 18, 2017
http://dx.doi.org/10.1287/inte.2016.0860
Copyright: © 2017 INFORMS

Abstract. The New York City Police Department (NYPD), the largest state or local police force in the United States, is charged with securing New York City from crime and terrorism. The NYPD’s Domain Awareness System (DAS) is a citywide network of sensors, databases, devices, software, and infrastructure that informs decision making by delivering analytics and tailored information to officers’ smartphones and precinct desktops. DAS development began in earnest in 2008; since then, the NYPD has used the system to employ a unique combination of analytics and information technology, including pattern recognition, machine learning, and data visualization. DAS is used throughout the NYPD, and the DAS software has been sold to other agencies, bringing in revenue for New York City. Through improving the efficiency of the NYPD’s staff, DAS has generated estimated savings of $50 million per year. Most importantly, the NYPD has used it to combat terrorism and improve its crime-fighting effectiveness. Since DAS was deployed department-wide in 2013, the overall crime index in the city has fallen by six percent.

Keywords: analytics • law enforcement • public service • predictive policing • counterterrorism • data visualization • pattern recognition • crime prevention • machine learning

With 8.5 million residents, New York (“the City”) is the largest city by population in the United States (U.S. Census Bureau 2015); in 2014, it was visited by approximately 56 million tourists (New York City Mayor’s Office 2015). New York is the world’s most economically powerful city (Florida 2012) and the center of the American financial industry, with a gross metropolitan product of $1.6 trillion (U.S. Bureau of Economic Analysis 2015). It contains uniquely significant cultural, historical, and governmental sites (e.g., the World Trade Center, the United Nations, and Times Square).

The New York City Police Department (NYPD) is charged with securing the City, its inhabitants, and its visitors from both crime and terrorism. To accomplish this mission, the NYPD employs 36,000 uniformed officers, constituting the largest state or local police force in the United States (U.S. Bureau of Justice Statistics 2011). Between 1993 and 2015, the City’s crime rate fell by roughly 75 percent (NYPD 2016a); part of this decline can be attributed to changes in police tactics and improvements in law-enforcement analytics (Zimring 2012). Concurrent with this crime decrease, the City suffered a devastating act of terrorism on September 11, 2001. The NYPD again responded by innovating; the department created a Counterterrorism Bureau, the first of its kind at a U.S. municipal police department (Dickey 2009).

The NYPD has a history of applying analytics to make officers more effective and efficient. For example, the NYPD’s structured process for examining crime statistics, known as CompStat, was developed in 1993 by Jack Maple and Commissioner William J. Bratton, and has since been the foremost example of law-enforcement analytics (Henry 2002). The purpose of CompStat is to ensure the accountability of precinct commanding officers, improve performance, and share information (Silverman 1999). CompStat meetings occur weekly and feature direct questioning
of precinct commanding officers by the department’s executive staff, informed by statistical analysis. CompStat began at the NYPD, but similar processes inspired by its methodology have since been adopted by other police and governmental agencies around the world (Weisburd et al. 2003).

Complementing the CompStat process, the Domain Awareness System (DAS) is the NYPD’s new means of employing analytics and operations research to inform officers’ decisions. DAS is a network of sensors, databases, devices, software, and infrastructure that delivers tailored information and analytics to mobile devices and precinct desktops. The system is now deployed across every police precinct in the City, on all officers’ smartphones, and on the tablets in all police vehicles. To the best of our knowledge, no police department in the world shares information and delivers analysis to its officers as effectively.

Problem Formulation
Before the DAS deployment, the NYPD failed to deliver the right analytics to the right place at the right time. As an organization, the department had accumulated a tremendous amount of information but had limited means of sharing it among its officers. Much of its information was available only to officers in the precinct house with permission to access stand-alone siloed software applications, and there were little analytics, operations research, or data visualization techniques applied to give the officers any insight. The NYPD’s situational awareness was also impaired by a lack of real-time sensors throughout the City, to say nothing of analytical tools that could make that sensor data actionable. Resource allocation and crime analysis had evolved little since the development of CompStat. Officers’ decision making was hampered by a failure to take full advantage of the information at hand, and the service the NYPD provided to the public suffered as a result.

NYPD officers’ decisions can be divided into two categories: tactical and strategic. Tactical decisions typically must be made quickly, sometimes in a matter of seconds, and usually pertain to a single incident. Strategic decisions are wider in scope, involve resource allocation or crime analysis, and are typically made by the commanding officer of a precinct under less time pressure. Below, we consider each of these types of decisions in detail, and discuss the problems the NYPD experienced in making these decisions before it deployed DAS.

As an example of a tactical decision, consider a 911 emergency phone call (known in NYPD lingo as a job) for a domestic disturbance. The NYPD’s method of directing officers to jobs had changed little in the 50-plus years since police radios were deployed. First, the caller speaks to an operator and describes the emergency; the operator’s transcription of the call is then passed to a dispatcher. The dispatcher assigns the job to officers in the vicinity, stating over the radio channel that the job is a domestic disturbance and reading out the address. The radio dispatch channel is typically congested (and can sometimes even become backlogged with a prioritized queue of jobs); therefore, the responding officers are usually provided little, if any, additional information. When they arrive at the location, however, these officers will need to decide very quickly how to manage the situation at hand.

Consider how much additional relevant information the NYPD likely had about this job. Perhaps a warrant was in place to arrest a person who listed this building as his (her) home address. Unless the responding officer had memorized the thousands of people in the NYPD’s warrants database, that officer had no way to know about the warrant. Perhaps, because of a history of domestic violence at this address, this job is the eighth call in the past 10 months. Only the precinct’s domestic-violence specialist, who has access to the database of domestic-violence records, would have this information, and the specialist would have to be sitting at his (her) desk to find it. The responding general patrol officer usually had no knowledge of any prior domestic incidents. In essence, the responding officers were forced to manage each situation, literally making life-and-death decisions, without any historical context. This is a big data problem tailor-made for analytics.

As another example of tactical decision making, consider an NYPD detective working on a terrorism case. The suspect is known to drive a vehicle with a particular license plate. The detective would like to interdict the vehicle and arrest the suspect; however, because the detective has no awareness of where in the City that vehicle is located, or where it will be in the future, he (she) must wait for the vehicle to appear again, perhaps while the suspect is en route to an attack. This is
a situation where predictive analytics focused on the vehicle’s driving patterns would be extremely helpful.

These types of problems also extend to the NYPD’s strategic decision making. Imagine a precinct’s commanding officer determining where to place units on patrol. Precinct commanders historically have made these decisions using hot-spot maps on which every crime in the past 28 days is marked by location. The precinct commander would simply circle clusters by eye, declare them as hot spots, and assign units to patrol at those locations; this is known as “putting the cops on the dots.” Hot-spot maps were made using a time-consuming process that did not allow easy access to the underlying records. Many analysts disliked the software used in this process so much that they chose to use handmade push-pin maps to visualize crime. Predictive analytics could greatly improve this state of affairs, allowing the commanding officer to use the NYPD’s data to more effectively deploy officers in specific locations.

Additionally, the strategic decision making that occurred at the weekly CompStat meeting was based upon a static table of numbers. Figure 1 is an example of one of these static tables—complaints are counted

![Figure 1](http://example.com/image.jpg)

**Figure 1.** (Color online) This CompStat page summarizes crime statistics for the 41st precinct, a neighborhood in the South Bronx, near the start of 2016. Without any data visualization, this static table is difficult to interpret.
starting with the seven major felony categories (murder, rape, robbery, felony assault, burglary, grand larceny, and grand larceny of an automobile) for three periods: week to date (i.e., the prior seven days), 28 day, and year to date. The rows in the figure display the number of complaints in the major-felony category (seven majors) occurring in transit, the number of complaints occurring in housing developments, three misdemeanor crimes of interest, and the number of shooting victims and incidents. The only analysis presented on this sheet is a year-to-year percentage change with very little data visualization applied. For anyone who has not spent years studying similar pages, picking out the most important trends is very difficult. Since a static page does not give access to the underlying records, using it to pick out geographic clusters or potential patterns is also impossible. The situation is ripe for modern data visualization.

Project History
The 9/11 Commission emphasized the importance of information sharing at all levels of government and across organizations (National Commission on Terrorist Attacks Upon the United States 2004). To this end, the DAS project began in earnest in 2008 in the NYPD’s Counterterrorism Bureau as a central repository for sensor data. The first sensors included in our system were private-sector security cameras around high-risk sites in lower Manhattan; we then expanded our sensors to include license plate readers (LPRs) and radiation sensors. In 2010, we began to integrate geocoded NYPD records, such as complaints, arrests, 911 calls, and warrants, to give context to the sensor data. At that point, we realized that our officers were being overwhelmed by the amount of data they were being asked to process. We then built alerting and pattern-recognition analytics into DAS to help officers manage the data and make better-informed counterterrorism decisions.

In 2013 we recognized the DAS software’s utility for general policing, in addition to counterterrorism purposes, and deployed DAS to every precinct in New York City. This rollout necessitated the construction of a robust network infrastructure to replace the outdated commercial networks on which we had been relying. As part of this expansion, we developed and deployed predictive policing algorithms to aid in resource-allocation decisions and a CompStat 2.0 data visualization package to aid in crime analysis.

In late 2014, recognizing that some of the policing decisions with the greatest impact are made in the field rather than in the precinct house, we developed and deployed a mobile version of DAS, which we optimized for smartphones and tablets. The production mobile deployment began in June 2015 and was completed in April 2016; all 36,000 NYPD officers now have a smartphone with access to DAS.

The DAS project has primarily been funded by approximately $350 million of U.S. federal homeland security grants. Our mobile deployment was made possible by a $140 million grant of federal forfeiture funds, drawn primarily from the fine on the bank BNP Paribas for violating international sanctions on Iran, Sudan, and Cuba (New York City Mayor’s Office 2014). The total investment in the program over the seven-year period since the project began is approximately $600 million.

Technical Solution
INFORMS defines analytics as the scientific process of transforming data into insight for making better decisions. DAS employs three analytic methods to inform tactical decision making (sensor alerting, automated pattern recognition, and real-time 911 response analytics), and two analytic methods to inform strategic decision making (predictive policing and CompStat 2.0). Each analytic method is built into the DAS software, which delivers the analytics directly to our officers. Next, we discuss each of these methods.

Sensor Alerting
The amount of information in DAS constitutes a big data problem. Audio gunshot detectors, 911 calls, license plate readers, closed-circuit television cameras, and environmental sensors all feed real-time data into DAS. As of April 2016, it also contains the following records: two billion readings from license plates (with photos), 100 million summonses, 54 million 911 calls, 15 million complaints, 12 million detective reports, 11 million arrests, two million warrants, and 30 days of video from 9,000 cameras. All of these sensors and records are accessible easily from the DAS interface. Text records are searchable via an indexed search engine.
Figure 2. The NYPD’s three-step alerting process, which we show in this diagram, allows us to take data from our sensors and use these data to inform law enforcement decision making.

The NYPD applies prescriptive analytics to each sensor feed; any reading that triggers an alarm is subject to a series of standard operating procedures, summarized as a three-step process. First, as we will explain, DAS triggers an automated alarm. Next, an officer examines an alarm and adjudicates it in the DAS software, informed by the larger context in DAS, such as other sensor feeds and all records geocoded in the vicinity of the alarming sensor. The adjudication step prevents the NYPD from deploying resources to a false-positive alarm. Finally, if the officer judges the alarm to be legitimate, a police response follows. Figure 2 graphically shows this process, and Levine and Tisch (2014) describe it in more detail. Each sensor feed has its own tailored mechanism for triggering an automated alarm; next, we discuss each of them in turn.

The 911 feed uses a logic-based rule set. For example, for counterterrorism purposes, any 911 call for shots fired near a high-risk building (e.g., the Empire State Building or World Trade Center) triggers an automated alarm; this constitutes a filter based on call type and location. Additionally, officers can customize their own alarms for jobs related to their area of responsibility and mission set; for example, the supervisory lieutenant in the 40th precinct can receive an alarm for every 911 call in that precinct for situational awareness.

The NYPD has deployed a set of gunshot detectors, widely referred to as ShotSpotters, in select areas of the City. These detectors are audio sensors that can triangulate the location of gunshots using acoustic correlation processing (Showen 1997). The technology also uses human-in-the-loop error checking; a trained expert listens to each audio recording before the alarm is generated to weed out false positives (e.g., caused by fireworks and truck backfires).

As part of the DAS program, we have deployed a network of LPRs in various locations around the City. We term the data the LPRs enter into DAS as “reads.” Currently, the NYPD deploys approximately 250 mobile detectors, which are mounted on police vehicles, and 50 fixed detectors that cover every lane of traffic on every bridge and tunnel into and out of Manhattan, and a line across Canal Street that fences off lower Manhattan. The LPRs use optical character recognition (OCR) to extract the alphanumeric characters from each plate (Anagnostopoulos et al. 2006); this is necessary because the system ingests three million reads per day, far more than can be examined manually. DAS compares the reads to terrorist and criminal watch lists, and any match triggers an automated alarm to the appropriate NYPD personnel.

DAS incorporates approximately 9,000 networked closed-circuit television cameras, which are owned and operated by a combination of public- and private-sector entities. The NYPD has deployed video analytics on a subset of these cameras to trigger automated alerts.
when one of these cameras detects suspicious behavior, including abandoned packages or movement in prohibited areas. These video analytics algorithms use background subtraction to compare the current image to a time-averaged one, and are tuned to generate an alert when an object the size and shape of a package is motionless for a period of time. Details of these algorithms are available in Bayona et al. (2009).

Environmental sensors are the final sensor type on which DAS generates alarms. These include a large set of radiation sensors, such as mobile detectors worn on the belts of police officers and fixed sensors deployed on the roofs of police facilities, and a smaller number of chemical sensors. Automated alarms on these sensor feeds are both logic based and threshold based. For example, some sensors are configured to generate an alarm whenever they detect radiation levels higher than a specific threshold. Other radiation sensors are configured to generate an alarm anytime they detect a particular radioisotope, such as uranium or plutonium, regardless of the radiation level.

**Automated Pattern Recognition**

We have implemented rules-based pattern recognition on our fixed LPR data streams. These predictive-analytics algorithms inform officers where and when a particular license plate of interest is likely to be scanned in the near future. Our custom algorithms search for two types of patterns: time and place, and routing. Pattern recognition is performed only on vehicles on an NYPD watch list; the vast majority of driving patterns are not relevant to law enforcement. Additionally, scans from watch-list vehicles are inspected for misreads during the alerting process described earlier; thus, the data entering into this process have already been cleaned.

Time-and-place patterns occur when a vehicle passes the same LPR on multiple occasions on a particular day of the week and at a particular time. For example, a vehicle that crosses the Brooklyn Bridge inbound around 9 AM every Tuesday has a time-and-place pattern. To detect such a pattern, the algorithm parses the recent history of a plate on the watch list and counts the number of times that plate was scanned by each fixed LPR in each of the 24×7 hour-long blocks in a week. Plates that are scanned frequently in the same block by the same LPR are considered to be in a time-and-place pattern.

A routing pattern occurs when a vehicle passes a fixed LPR and then some time later passes a second fixed LPR. For example, if a vehicle entered the City via the Queensboro Bridge and then traveled southbound on the FDR Drive, it would pass two of our LPRs in a 15-minute time span (depending on traffic). If the vehicle traveled this route once on Tuesday morning, and then again on the next Monday afternoon, that would constitute a routing pattern. To detect a routing pattern, the algorithm parses the history of a plate on the watch list and counts the sequential pairs of fixed readers that have scanned the plate. Sequential pairs that exceed a predetermined threshold are considered to be a routing pattern. We put no limit on the amount of time that can separate a sequential pair, because we want to detect round trips of arbitrary duration (e.g., a vehicle that leaves the City over the George Washington Bridge on a Monday morning and returns on Friday afternoon).

Pattern-recognition predictive analytics are useful for our officers because they forecast the future locations of watch-listed vehicles without requiring the officer to study the scanned history of a plate; of course, that history is available in DAS if the officer would like to inspect it. Forecasting the future locations of vehicles makes interdictions much easier, eliminating the need for officers to constantly monitor the vehicle, which is expensive in terms of staffing, or place a GPS device on the vehicle, which requires access to the vehicle. An officer who wants to interdict a vehicle with a time-and-place pattern would simply wait at the time and place that the vehicle is forecast to appear. An officer who wants to interdict a vehicle with a routing pattern would wait to receive an alert that the vehicle has passed the first LPR in the pattern. That officer would then quickly travel to the second LPR in the pattern and await the arrival of the vehicle of interest.

The NYPD has also implemented simple pattern detection to find vehicles that caravan together. For example, if a vehicle on our watch list frequently travels in close proximity to a second vehicle, then that information is likely to be relevant for investigators. Each time the watch-listed vehicle is observed by a fixed LPR, the algorithm records other vehicles that passed the LPR shortly before and after that vehicle. Vehicles that appear on this caravan list with a frequency
higher than a chosen threshold are considered possible members of a caravan and subject to additional investigation.

**Real-Time 911 Response Analytics**

The third tactical decision-making technical solution we discuss is tailored for officers responding to 911 calls. In the *Problem Formulation* section, we describe the example of an officer who is responding to a domestic-violence call but is given only the address and type of emergency by the radio dispatcher. Now, through the DAS software deployed on officers’ smartphones, the responding officer has access to the billions of records stored in DAS. We have implemented a series of rules to solve this big data problem and prioritize the most relevant records.

When an officer is assigned a job, his (her) smartphone receives an alert via a text message-like notification. The officer can then directly view the caller’s statements, as transcribed by the operator. Officers in the field had never previously been able to view this information—that is, the problem described in the caller’s own words. The officer can also directly contact the caller via phone if further clarification is necessary. Prior to DAS, all contact with the caller took place through the dispatcher and operator; our application has eliminated the middleman.

Historical information is also extremely useful for the responding officer; in the DAS mobile application, the responding officer can view a history of the job’s address, as drawn from NYPD records. The NYPD uses analytics to whittle down the billions of records stored in DAS to a select few that can be viewed in the short amount of time before the officer arrives on the scene. Any records that have a matching address will be displayed in a prioritized order, with records indicating the potential for violence (e.g., active warrants, previous calls for shots fired, and records of emotionally disturbed persons) brought to the top. Some records have associated photos; these are displayed alongside the text of the record. Officers can also do radius searches around the job’s geocoded location for more extensive situational awareness.

NYPD patrol officers are deployed in pairs. Our policy states that the officer in the passenger seat should review and read aloud relevant information in DAS, while the officer operating the vehicle drives as normal to the job’s location.

For additional situational and supervisory awareness, we also allow officers to view information about jobs that have been assigned to other NYPD officers. Supervisors can read about the jobs to which their subordinates are responding, and if a job appears to warrant additional attention, the supervisor can also respond before the assigned officer even knows that backup might be necessary.

**Predictive Policing**

The first technical solution for strategic decision making we describe is predictive policing. As we describe in the *Problem Formulation* section, resource-allocation decisions at the NYPD have traditionally been informed with hot-spot maps, sometimes still constructed by hand. The purpose of the NYPD’s predictive policing algorithms is to better inform these decisions with predictive analytics. For example, imagine that a precinct has been experiencing a problematic rise in robberies over the past month. Where should the commanding officer deploy his (her) officers to prevent additional robberies from occurring?

We have deployed predictive policing algorithms for six types of events: shootings, burglaries, felony assaults, grand larcenies, grand larcenies of motor vehicles, and robberies. Each type of event has an individually tuned and independent algorithm. The appendix provides modeling details on these algorithms. For each crime type, the algorithms highlight 10 boxes, each of which measures 500 feet by 500 feet, per precinct to aid precinct commanders’ resource-allocation decisions.

Table 1 shows the success rates of these algorithms during a 24-week cross-validation period compared to traditional 28-day hot spots. A crime is counted as successfully predicted if it occurred in one of the 10 boxes highlighted for that period. That is, 3.2 percent of burglaries during the 24-week cross-validation period fell in boxes highlighted by 28-day hot spots; in comparison, 7.2 percent of these burglaries fell in boxes highlighted by the NYPD’s predictive algorithm. Indeed, for each crime type, we show a significant improvement over the 28-day hot-spot traditional technique. For shootings, the NYPD’s predictive algorithm outperformed the hot-spot technique by a factor of more than five.

To deploy a new analytical technique, such as predictive policing, we have found that taking a two-pronged
Table 1. We calculated the percentage accuracy of NYPD’s predictive algorithm vs. the traditional method during a 24-week cross-validation period.

<table>
<thead>
<tr>
<th></th>
<th>Burglary (%)</th>
<th>Felony assault (%)</th>
<th>Grand larceny (%)</th>
<th>Grand larceny (motor vehicle) (%)</th>
<th>Robbery (%)</th>
<th>Shooting (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional (28-day hot spot)</td>
<td>3.2</td>
<td>7.5</td>
<td>11.2</td>
<td>3.1</td>
<td>6.3</td>
<td>3.7</td>
</tr>
<tr>
<td>NYPD predictive</td>
<td>7.2</td>
<td>14.9</td>
<td>20.2</td>
<td>5.7</td>
<td>13.4</td>
<td>20.4</td>
</tr>
</tbody>
</table>

approach is helpful. First, create a more efficient version of the traditional technique; then develop a version of the new technique that has a user interface similar to the traditional technique. To implement predictive policing, we first built a more efficient version of hot-spot maps into DAS and made it available for any officer who wants to use it. Then, in our design of predictive policing, we built a user interface into DAS with controls that were nearly identical to those in the hot-spot interface. This eased the transition from the old technique to the new.

CompStat 2.0

The second technical solution for strategic decision making is CompStat 2.0 for crime analysis. As we describe in the Problem Formulation section, most of the discussion at the weekly CompStat meeting was framed around a static book of numbers that was compiled and published every week. However, decision makers find it difficult to pull relevant information from a table; graphs and other data graphics are a more natural way to investigate trends and patterns (Tufte 2001). We have modernized the CompStat process, now called CompStat 2.0, and turned the book into an interactive software module that allows officers to easily explore the statistics and see trends from within DAS.

CompStat 2.0 presents every number in the static CompStat book in an interactive form (Figure 3). If a user clicks on a number, DAS brings up all the records included in that number and marks them on the map. We have also constructed a data visualization engine to enable the user to explore trends and patterns in the statistics. Bar and pie charts of categorical data and line charts of temporal data are available with the press of a button.

In the past, analytic and data visualization expertise was concentrated at headquarters, and precinct commanders were at a disadvantage because they could not examine crime data with the same sophistication as headquarters staff. These new capabilities allow commanding officers and their crime analysts to detect spikes and patterns as easily as the most experienced analysts at headquarters can. Determining the effects of any innovative crime-fighting strategies is more straightforward. DAS has essentially democratized data visualization expertise; no longer are spreadsheet and slide-crafting experts necessary for commanding officers to understand the crime in their precincts.

Challenges

Police officers are a difficult audience to impress; some experienced officers have decades on the job and have seen many new technologies falsely promise to revolutionize policing. The NYPD’s 170-year history includes many examples of such projects. However, unlike these other technologies, DAS was designed for cops by cops, and that has been the secret to our success.

We involve our officers at every step of our development and deployment processes. In addition to helping us build better products, working side by side with our officers allows us to use cops as our evangelizers and trainers in the precinct houses, rather than having to rely on civilian contractors to train NYPD personnel. Law enforcement can be an insular community, and the NYPD is no exception. A fellow officer explaining the benefits of a new technology is far more convincing than a civilian who has never been on patrol.

NYPD officers also tend to be skeptical of analytics solutions that run counter to their instincts. That is, if we deployed a black-box solution, our officers would likely not trust or use the outputs. For this reason, we make sure to employ techniques for which the underlying records influencing the outputs can be made visible, thus enabling officers to see the algorithm’s drivers. For example, in our predictive policing tool, when the algorithm predicts a box as relatively high risk for a specific crime type, the officer can
At various steps in its evolution, the DAS project did face opposition. Organizationally, the NYPD’s information technology bureau (ITB) has traditionally been more concerned with keeping the existing outdated systems operational than with innovative policing. In the early years of DAS, our team was unable to get much cooperation and support from this bureau. ITB’s attitude was “it wasn’t invented here.” However, when Police Commissioner William J. Bratton was appointed in 2014, he promoted the head of the DAS project to lead ITB. DAS has since become the primary vehicle enabling the NYPD to modernize its data management and decision making.

DAS has also faced some opposition from civil liberties groups concerned with police surveillance of public spaces. At the start of our project, we thought deeply about these issues and preemptively wrote a privacy policy that governs the NYPD’s use of the information in DAS. A draft of this policy was released for public comment, edited, and published in its final form in April 2009 (NYPD 2009).

**Impacts**

**Adoption by the NYPD**

We began deploying DAS on desktops across the NYPD in 2013; since then, we have used several metrics to measure its adoption and its constituent components. Two of those metrics are the number of unique individuals that used the software each week, called unique logins, and the number of record searches they executed, simply called searches. Figure 4 depicts the DAS adoption throughout the desktop deployment and beyond; the number of weekly unique logins grew by a factor of 10, and the number of searches grew by a factor of 100. In the mobile environment, we use those same metrics, separating tablet and smartphone logins; we also track the number of job clicks—the count of 911 calls where an officer explores any of the additional information provided. We plot these metrics in Figure 5, which demonstrates the wide usage of the DAS...
mobile application. At the end of 2015, we hit 100,000 mobile job clicks per week. We believe this demonstrates the officers’ embrace of the new technology and analytics.

**Public Safety**

The DAS project is critically important for the people of New York City. Every day, the NYPD uses the analytics built into DAS to save lives, arrest criminals, prevent crimes, and keep the public and our officers safe. In this section, we discuss evidence for these public-safety impacts.

The Counterterrorism Bureau has been using DAS to secure the City since 2008. For example, when the terrorists who bombed the Boston Marathon in April 2013 were rumored to be traveling to New York City, we entered their vehicle’s plate into our LPR watch list to alert our officers if the vehicle was observed. This triggered our OCR algorithms to check every scanned license plate, which now number approximately three million per day across both mobile and fixed detectors. Fortunately, the terrorists were interdicted before they reached New York City; however, if a match had been detected and an alert issued,
our officers would have responded immediately and intercepted the vehicle. Additionally, our video analytics alerting has detected several abandoned packages deemed suspicious enough to warrant physical inspection by the NYPD’s bomb squad wearing full protective blast suits; luckily, these packages all turned out to be false alarms.

DAS alerting has also had a strong impact on traditional crime fighting. For example, on December 1, 2015, Diomedes Valenzuela abducted his three-year-old daughter after assaulting his wife, and fled in his minivan. Officers entered the vehicle’s plate into a watch list, again causing the OCR algorithms to check every scanned plate for a match. These algorithms promptly triggered an alarm when the vehicle crossed the George Washington Bridge heading into New Jersey. This information was passed to pursuing officers, who quickly found and apprehended the perpetrator and prevented any further harm to the young girl.

Numerous shootings have been detected with both the audio gunshot detectors and the DAS video camera network. In the areas entered by gunshot detectors, our analyses have shown that 75 percent of shootings are not called in to 911 (New York City Mayor’s Office 2016); hence, on many occasions, the gunshot alerts in DAS that are triggered by acoustic correlation algorithms are our only signal that a shooting has occurred. These alerts have resulted in improved police responses, additional evidence collected, and more violent criminals being arrested and charged.

The LPR pattern-detection algorithms have also been used to catch criminals. For example, a person with an open warrant for wire fraud, and whose vehicle was on an NYPD watch list, was found to have a time-and-place pattern entering Manhattan via the Brooklyn Bridge. The NYPD dispatched a police car to the Brooklyn Bridge at the predicted date and time; the officers in the car spotted the suspect’s vehicle as expected, pulled him over for a traffic violation, and then arrested him.

Although the new mobile technology was only deployed in 2015, NYPD officers are already using it to provide better service to the public and save lives. For example, in January 2016, officers from the 19th precinct were investigating a string of eight bank robberies (NYPD 2016b). As part of this investigation, they configured the rules-based alerting on their smartphones to notify them whenever a 911 call pertaining to an active bank alarm was received. When they received such a call, DAS triggered alerts on their smartphones. These officers responded before they were notified by radio dispatch and arrested the suspect, who then confessed to the other robberies in the pattern. Since the radio dispatch channel was congested at the time of the call, the officers would likely not have been able to respond quickly enough to make the arrest without DAS analytics. Access to wanted flyers immediately upon creation (Prendergast 2015), direct communication with 911 callers (Fleischer 2015), and searchable department records (Valenti 2015, Parascandola 2015) have all also resulted in the arrest of criminals. The events covered in the news are just the tip of the iceberg; we have collected dozens of success stories for DAS mobile analytics.

Additionally, predictive policing is having an impact on our deployments. For example, Sergeant Joseph Klubnick of the 83rd precinct in the Bushwick neighborhood of Brooklyn used our predictive policing algorithm to identify where to place an anti-burglary team. The algorithm highlighted a region approximately three square blocks in size. After two hours of surveillance, the team observed and apprehended an individual actively engaged in burglarizing a basement in that region; the individual was charged with burglary in the third degree. Without predictive policing, this team could have been placed somewhere else in the precinct and not made the arrest.

Finally, the CompStat 2.0 software module is now the data visualization tool used in the NYPD’s weekly CompStat meeting, which is attended by the NYPD’s highest-ranking executives and hundreds of supervisors; at this meeting, the analytics provided by DAS inform decisions on crime-fighting strategies citywide. Officers use the software’s analytics during the meeting to visualize and investigate many of the topics discussed. This would never have been possible with the static books, with which all analyses must be finalized before the meeting starts.

Since DAS was deployed department wide in 2013, it has resulted in the tremendous growth in usage shown in Figure 4, and the overall index crime in the city has fallen by six percent. This is driven by the 10,000 fewer burglaries, robberies, and grand larcenies that occurred during the two-year period. Although we cannot give DAS credit for the entirety of this decrease,
all of this anecdotal evidence, along with officer testimonial, lead us to believe that it is having a significant impact and making our officers more effective and efficient.

Financial

DAS has had a significant financial impact in NYPD staffing. Before DAS, the NYPD used 500 of its officers as crime analysts; four analysts staffed a typical precinct, and groups of analysts were at the borough and city levels. By making these analysts more efficient, we have been able to assign some of these officers back to patrol. Additionally, we have made the Detective Bureau’s investigations more efficient, an improvement that affects 5,000 officers. To approximate the total savings, we use an estimate of $100,000 annual salary and benefits per officer, a 50 percent increase in efficiency for crime analysis, and a five percent increase in efficiency for investigations. We estimate savings from staffing to be approximately $50 million per year. We have also brought in an additional $1 million from external sources, as we describe later.

Endorsements

In our Edelman presentation (INFORMS 2016), we quote Police Commissioner William J. Bratton: DAS “helps us deal with the traditional areas of preventing crime and disorder, but it is also crucial to our 21st-century mission: the challenges of terrorism.” He put DAS front and center in his plan of action, stating that “making data accessible to the field can significantly improve the Department’s effectiveness” (NYPD 2015).

Officers on the street are also supportive of DAS (INFORMS 2016). Officer Robert Guglielmo of the 19th precinct states “everybody on my team, we all use it… it gives us a step up on the perpetrators.” Dan Higgins, a detective in the Counterterrorism Bureau, says “the DAS is absolutely a significant crime fighting tool for the NYPD.” Michael Fragedis, an officer in the 67th precinct, says “it’s taken us to the next level… it’s got my thumbs up.”

Transportability and Adoption by Other Agencies

The NYPD has provided demonstrations of DAS to hundreds of law-enforcement and public-safety agencies around the world. Many of these agencies have taken ideas from the demonstrations and built some of the NYPD’s best practices into their own platforms. Others, such as the Washington D.C. Metro Police, the Singapore Police Force, and the Brazilian National Police have purchased the DAS software from Microsoft, our software developer, and have used it to secure high-profile governmental and cultural sites, the 2014 World Cup, and the 2016 Summer Olympics. Microsoft has agreed to give New York City 30 percent of the revenue it derives from selling the software to other jurisdictions (Parascandola and Moore 2012); these purchases have already generated over $1 million of additional revenue for the City.

By openly publishing the details of our analytics and operations research techniques, we are providing other public-safety agencies with the tools and methods to adopt and incorporate these ideas into their own jurisdictions. In this way, the public benefits New York City has enjoyed from using these techniques will start to appear in other cities around the world.

Future Plans

We have a number of mechanisms to collect feedback from patrol officers, supervisors, and precinct commanders, including focus groups and email suggestion boxes. These sources, together with suggestions from NYPD analytics professionals and our colleagues at INFORMS, have generated new ideas for the next generation of DAS.

Since we deploy new versions of the DAS software several times a year, DAS is a continually evolving system and new features will enter production over time. One enhancement will be automated pattern detection for crimes using supervised machine-learning techniques (Wang et al. 2013). This effort will simplify the existing manual and memory-based process, which does not involve DAS. Additionally, we have recently installed GPS devices in all our marked vehicles; the data set that results from these devices will have implications for optimization of patrol locations and 911 response times.

Acknowledgments

The authors thank the men and women of the NYPD. The authors have benefitted from the many ideas of Police Commissioner William J. Bratton, Chief of Department James O’Neill, Deputy Commissioner John Miller, Chief of Counterterrorism James Waters, Inspector Christine
Appendix

To discuss the mathematical details of these predictive policing algorithms, some notation is required (Table A.1).

The six predictive policing algorithms have a similar structure. They all begin by dividing the area of interest, such as a commanding officer’s precinct, into a regular grid (e.g., boxes measuring 500 feet by 500 feet). For each box in the grid (indexed by \( i \)), the algorithm calculates a score, \( S_i \), by summing the weights \( w_{i,j} \) associated with each of the relevant records in that box:

\[
S_i = \sum_j w_{i,j}.
\]

The sum runs over the relevant records related to events that occurred in the box, indexed by \( j \). After the algorithm calculates the scores, the boxes with the highest scores are highlighted on a map in the DAS user interface and are used to recommend the deployment of police resources.

Each of the six predictive policing algorithms incorporates a different method of calculating the weight associated with each relevant record. As an example, let us examine the algorithm used for predicting shootings, which uses two types of relevant records: 911 calls for shots fired and the NYPD’s shootings database. The fields in each record influence the algorithm are calculated as

\[
w_{i,j} = \left( a_i + \left( \frac{1.095 - d_j}{1.094} \right) \left( 0.5 + a_2 \cos \left( \frac{2\pi d_j}{365} \right) \right) \right) c_{i,j},
\]

where \( a_i \) and \( a_2 \) are seven constant parameters tuned via gradient descent. This weight is constructed from three terms multiplied together. The first term is a linear function that decreases in value the further the record has occurred in the past. The algorithm only includes records of events that occurred within the last three years; hence, we use 1,095 (365 days \( \times \) 3) in the fraction’s numerator. The second term accounts for seasonal variations in crime; see U.S. Bureau of Justice Statistics (2014) for more detail on this phenomenon. This function is sinusoidal, with a period of 365 days; therefore, it peaks during the same season in previous years. The last term \( c_{i,j} \) is calculated using a combination of the record’s characteristics:

\[
c_{i,j} = \begin{cases} 
1 + a_3 v_{\text{dms}} + a_4 v_{\text{db}} + a_5 v_{\text{mot}} & \text{for shooting records}, \\
\left( a_6 \left( 1 + a_7 v_{\text{con}} \right) \right) & \text{for shots-fired 911 calls}.
\end{cases}
\]

The \( v \) terms are binary values dependent on each record’s characteristics, as described in Table 1.

The last pieces of information necessary to calculate the weights are the seven parameters, \( a_i \). These parameters are optimized using the machine-learning technique of gradient descent. We use the combination of parameters that maximizes the number of shootings that occur in the top-10 scoring boxes, training on 52 weeks of shooting trials (April 2014 to March 2015), and containing 1,196 shooting incidents. For each week, the algorithm looks at the previous three years of data and predicts 10 boxes for each precinct. Each shooting that occurs in a predicted box during the following week counts as a success, and each shooting that occurs outside of the predicted boxes is a failure. The algorithm then moves on to the next week in the data set. To verify that the algorithm has not been overtuned, we check its accuracy on 24 weeks of cross-validation data (April 2015 to September 2015), which contain 673 shooting incidents. The accuracy of the model on the training data is only slightly better than the accuracy on the cross-validation data (21.0 percent versus 20.4 percent), indicating that the model has not been overtuned.

Table A.1. This table summarizes the mathematical notation used in the predictive policing algorithm for shootings.

| \( i \) | Index corresponding to the box number in the precinct’s grid |
| \( S_i \) | Score of box with index \( i \) |
| \( j \) | Index corresponding to each relevant record geocoded in a box |
| \( w_{i,j} \) | Weight associated with the \( j \)th relevant record in box \( i \) |
| \( d_j \) | Days since the event described in record \( j \) occurred |
| \( c_{i,j} \) | Part of the weight that depends on the structured characteristics of record \( j \) in box \( i \) |
| \( v_{\text{dms}} \) | A binary value, which is 1 if the shooting is related to a domestic disturbance, and 0 otherwise |
| \( v_{\text{db}} \) | A binary value, which is 1 if the shooting is related to a social club, and 0 otherwise |
| \( v_{\text{mot}} \) | A binary value, which is 1 if the shooting’s primary motive is gang or revenge related, and 0 otherwise |
| \( v_{\text{con}} \) | A binary value, which is 1 if the 911 shots-fired record was confirmed with physical evidence, and 0 otherwise |
| \( a_1 \ldots a_7 \) | Seven constant parameters tuned via gradient descent |
gunshot detectors in a limited area of the City. Therefore, if we included gunshot detector records in our predictive algorithm, the calculation of the boxes’ scores would be biased toward areas with gunshot detectors.

We have similar algorithms for five other crime types.

References


E. S. Levine is the Director of Analytics at the NYPD’s Counterterrorism Bureau, and has led the integration of analytic and operations research techniques into the DAS since 2010. He is responsible for the NYPD’s efforts in predictive policing, pattern recognition, and data visualization. Prior to joining the NYPD, he was the Chief Scientist at the Department of Homeland Security’s Office of Risk Management.
and Analysis where he designed the first U.S. all-hazard homeland security quantitative risk assessment. He began his career in government as a Presidential Management Fellow and is a member of INFORMS, the Society for Risk Analysis, and the American Association for the Advancement of Science. He is the author of Applying Analytics: A Practical Introduction, a textbook for undergraduates, as well as numerous peer-reviewed journal articles. He holds a PhD in astrophysics from the University of California at Berkeley, a master’s in advanced study of mathematics from the University of Cambridge in England, and a bachelor’s in astronomy and physics from Harvard University.

Jessica Tisch is Chief Information Officer at the New York City Police Department (NYPD) and has led the development and deployment of the Domain Awareness System (DAS) since its inception in 2008. As head of the Information Technology Bureau, she is responsible for modernizing and maintaining the NYPD’s technology, including its network infrastructure, data storage, end-user devices, and communications equipment; additionally, she manages the 911 system for the entire City. Previously, she served as counsel to the Police Commissioner and the Counterterrorism Bureau’s Director of Policy and Planning, where she was the lead author on the NYPD’s book Engineering Security: Protective Design for High Risk Buildings. She is passionate about using technology and analytics to enable government to better serve the public. She holds a JD, MBA, and bachelor’s degree in government from Harvard University.

Anthony Tasso, Inspector, is the commanding officer of the NYPD Information Technology Bureau’s Strategic Technology Division and has worked on the Domain Awareness System since 2014. Previously, he served as commanding officer of the 83 (Bushwick) and 88 (Fort Greene) Precincts and has also served as executive officer in the 73 (Brownsville), 84 (Downtown Brooklyn), and 90 (Williamsburg) Precincts. As a precinct commander he successfully drove down crime and violence; for example, during his two-year command of the 83 Precinct, murder decreased 59% and shooting incidents fell 43%. In all of his roles he has focused on using new technology and analytics to improve police operations and keep the public safe. He began his career on patrol in Coney Island and now has 22 years of experience as a uniformed NYPD officer. He was also a first responder at the World Trade Center on September 11, 2001. He holds a bachelor’s in organizational science from St. Joseph’s College and is a 2009 graduate of the Police Management Institute at Columbia University.

Michael Joy, Captain, is the commanding officer of the NYPD Strategic Technology Division’s Operations Section and has worked on the Domain Awareness System since 2008. Previously he was the commanding officer of the Counterterrorism Bureau’s Project Management Office. In this role, he led the NYPD’s rapid deployment of mobile wireless security cameras, first used at the 2013 New York City Marathon, in response to the bombing of the Boston Marathon. He started his career on patrol in the subways in the Bronx and now has 14 years of experience as a uniformed NYPD officer. He holds a master’s from Boston University in computer information systems and a bachelor’s in sociology from SUNY Albany.