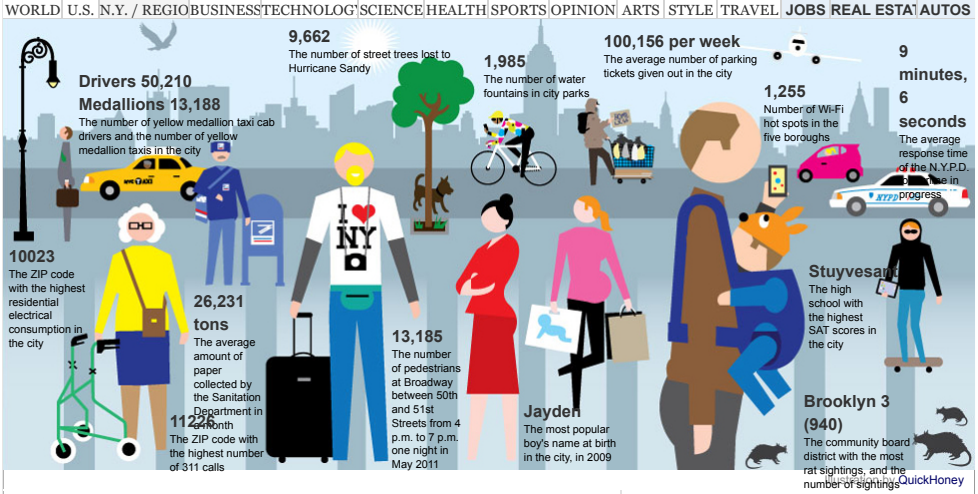
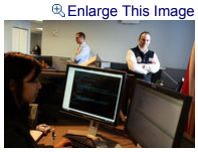


Illustration by QuickHoney

The Mayor's Geek Squad

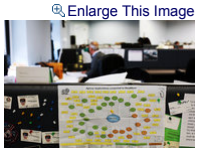
By ALAN FEUER
Published: March 23, 2013

It was a case for a digital Sherlock Holmes. Last fall, the city's Department of Environmental Protection wanted, finally, to crack down on restaurants that were illegally dumping cooking oil into sewers in their neighborhoods — congealed yellow grease is responsible, the department says, for more than half of New York's clogged drains. The question, of course, was how to find the culprits?



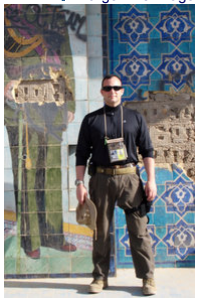
Todd Heisler/The New York Times

Michael Flowers, right, oversees a small group of tech-savvy and civic-minded statisticians working across from City Hall.



Todd Heisler/The New York Times

"All we do," Mr. Flowers said, is "process massive amounts of information and use it to do things more effectively."



Collection of Michael Flowers

Mr. Flowers worked in Iraq in 2005.

The antiquated answer would have been to have the health department send inspectors to restaurants on blocks with backed-up sewers and hope by chance to catch a busboy pouring the contents of a deep fryer into the street.

Enter the city's Office of Policy and Strategic Planning, a geek squad of civic-minded number-crunchers working from a pair of cluttered cubicles across from City Hall in the Municipal Building. They dug up data from the Business Integrity Commission, an obscure city agency that among other tasks certifies that all local restaurants have a carting service to haul away their grease. With a few quick calculations, comparing restaurants that did not have a carter with geo-spatial data on the sewers, the team was able to hand inspectors a list of statistically likely suspects.

The result: a 95 percent success rate in tracking down the dumpers. With nothing grander than public data, the Case of the Grease-Clogged Sewers was solved.

Data — or Big Data, as quantitative analysts will call it — is the tool du jour for tech-savvy companies that have realized that lurking in the vast pools of unprocessed information in their networks are solutions to some of today's most pressing and convoluted problems. A few years ago, Google, for example, took the 50 million most common keywords that Americans typed in search bars and tried to figure out, by comparing them with federal health statistics, where the H1N1 flu virus was to likely strike next.

According to a new book, "Big Data: A Revolution that Will Transform How We Live, Work and Think," the enormous quantity of information whirling through the ether can affect and enhance our quality of life. As the authors put it, "The change of scale has led to a change of state."

Now the city has brought this quantitative method to the exceedingly complicated machine that is New York. For the modest sum of $1 million, and at a moment when decreasing budgets have required increased efficiency, the in-house geek squad has over the last three years leveraged the power of computers to double the city's hit rate in finding stores selling bootleg cigarettes; sped the removal of trees destroyed by Hurricane Sandy; and helped steer overburdened housing inspectors — working with more than 20,000 options — directly to lawbreaking buildings where catastrophic fires were likeliest to occur.

"I think of us as the Get Stuff Done Folks," Michael Flowers who oversees the group, said. "All we do is take and process massive amounts of information and use it to do things more effectively."

Before being hired in 2009 by John Feinblatt, the mayor's chief policy adviser, Mr. Flowers didn't know much about computer code — let alone Bayesian statistics. From 1999 to 2003, he worked at the Manhattan district attorney's office, prosecuting homicides and drug crimes. When he left law enforcement, he moved to Washington, where he joined the power law firm Williams & Connolly and later took a job with the Senate Permanent Subcommittee on Investigations. Disenchanted by the smug homogeneity of Washington, Mr. Flowers leapt at the chance in 2005 to travel to Iraq with a team from the Justice Department to work on issues concerning mass graves and on Saddam Hussein's trial.

While serving in the Green Zone, Mr. Flowers was responsible for sending investigators to grave sites in the countryside and transporting witnesses against Mr. Hussein to his office — without getting either group blown up by roadside bombs. It came to his attention that military officers were using predictive informational techniques to determine where and when the bombs were likely to explode.

He borrowed those techniques when he returned to New York and went to work for Mr. Feinblatt with the initial, limited task of trying to understand in the early months of the recession what was causing mortgage fraud.

"We eventually realized there was enormous value in using all our data — together and proactively," Mr. Feinblatt said. "We'd already done the retroactive act of looking back for accountability's sake. So we tried to use the data prescriptively to figure out what might be coming next."

These days, Mr. Flowers, a relative amateur in data analytics, is the geek squad's chief tactician and resident asker of questions.

He allows the half-dozen post-collegiate techies working under him to ferret out the answers and, at age 43, he refers to them endearingly as "the kids." His office gives the impression of a high-tech start-up — but without the cool furniture. Nick O'Brien, 30 and the team's chief of staff, works standing at a lectern. Ben Dean, the 24-year-old chief analyst, sits on an ergonomic rubber ball.

One drawer in the filing cabinet is filled with spare neckties. These are to spruce team members up for meetings with "That Guy," as Mr. Flowers likes to call his boss, Mayor Michael R. Bloomberg, invariably chucking a thumb in the direction of City Hall.

Two weeks ago, with Mr. O'Brien in Texas for the South-by-Southwest conference, the rest of the team was working on a project to make the city's response to natural disasters like Hurricane Sandy more robust. Catherine Kwan, 24, was doing some "MacGyver stuff," as Mr. Flowers called it, correlating city information with data from utilities, like Con Edison, to put in place a system that would eventually detect, in real time, when a building's heat or lights were out.

"So what's the current ratio of Con Ed customer accounts per residential unit?" Mr. Flowers asked. (Ms. Kwan's answer: 0.88.) And the ratio of people per unit was "2.6," she said.

One of the benefits that come from working with the informational atoms of the city is an almost molecular understanding of New York itself. The youthful quants were surprised to learn, for instance, that it was mathematically possible to create safer streets by encouraging local businesses to keep their doors open later after dark. They also had not known that a significant percentage of 311 complaints derived from certain neighborhoods in Lower Manhattan — an area they now refer to jokingly as "whine country."

"What's impressed me most about this job," Mr. Flowers said, "is learning how insanely complicated this city is." He mentioned, in particular, the 900,000 buildings the city oversees and the 12,000 tons of trash it picks up daily.

"That activity is reflected in the data and on an amazingly detailed level. What we're really running here is an office of New Yorkology," he said.

WHAT THE CITY KNOWS about its 8 million residents is staggering. Contained in public archives is information about their boilers and their sprinkler systems, the state of their local taxes, the number of heart attacks and fires that occur inside their buildings and whether they have ever logged complaints about roaches or construction noise. Additional data is gathered about their businesses, their commuting habits and their children's test scores.

If a parking meter sits outside their apartment, the city knows how many cars have parked there on any given day, the number and dollar amount of tickets handed out and, of course, the identities of those who have received them.

"There's a deep, deep relationship between New Yorkers and their government," Mr. Flowers said, "and that relationship is captured in the data."

In all, a terabyte of raw information — enough to fill nearly 143 million printed pages — passes daily through Mr. Flowers's office, and his team's first job, he said, was to get that information into a comprehensible form: to, in effect, create a lingua franca for the bureaucracy's Tower of Babel. As Mr. Feinblatt put it, "The data will tell you a story, but only if you do certain things that encourages it to speak."

Among the first things the city did was to establish in 1989 the Commission on Public Information and Communication, or Copic, which under the aegis of the Public Advocate's office was charged with helping New Yorkers get better access to municipal information.

While "good-government" advocates like Noel Hidalgo, executive director of the Open New York Forum, which advocates for the use of technology in city management, have questioned the effectiveness of Copic, they have also said its existence laid the groundwork for the passage last year of Local Law 11, one of the country's most progressive open-data laws.

"Copic was Version 1.0 for greater transparency in public information," Mr. Hidalgo said, "but it needed updating for the 21st century. Now, with Law 11, there is the potential to radically change how government services are used by citizens. It opens the door to a unique partnership between the city and its residents so that people can come up with innovative ways of using information."

The law's chief provision created a clearinghouse called the Open Data Portal, which offers to the public hundreds of sets of city data, including the location of Wi-Fi hot spots, the results of restaurant inspections, yearly power use by ZIP code and maps of public parks.

Mr. Flowers is responsible for managing the portal — which just this month placed online all city data now available — but the information on the site has been employed by a wide variety of people. Two weeks ago, the Office of Financial Empowerment, a city agency that helps low-income residents, held a "hackathon" at which tech geeks using information from the portal built an internal scheduling system for the agency's counselors.

The portal has also served as a primary source for hackers in the city's Big Apps competition, which annually awards cash prizes to developers who have created applications like ones helping cyclists avoid city streets where accidents often occur and providing the locations of public restrooms.

With each passing week, it seems another hackathon — think hacking marathon, usually to a beneficial purpose — is mounted in New York. There was Foursquare's effort in January that resulted in the Nasdrunk app (it matched the closing value of the Nasdaq with smartphone check-ins at various city bars), and then there was Decoded Fashion, which took place during Fashion Week and was billed by its sponsor, Condé Nast, as the world's first fashion industry hackathon.

Though many of these events used proprietary data — Occupy Wall Street held a hackathon this month crunching numbers from its Hurricane Sandy relief effort — their diversity and frequency have created a kind of hothouse atmosphere, a local data frenzy in which private efforts at analysis have spurred city government on to do the same.

"I think New York is the natural place for Big Data," Mr. Flowers said. "We have the right culture. We have a mayor who understands that management is measurement. And, of course, we're big enough so that it makes sense analyzing the data that we have.

"All the pieces, all the structures, are in place," he said. "In New York, it's kind of like the triumph of the nerds."

ONE DAY THIS MONTH, Mr. Flowers, in a military swag vest from Iraq, was in his office kicking around ideas for future projects. Unlike the ascendant nerds he mentioned, he tends to speak in a soldier's clipped language: "Mission critical" or "Actionable outcomes."

This, indeed, was a "spitballing" session, and the plans being tossed around revealed where he would like to go next with his team. One idea was to analyze, and hopefully reduce, the time it takes for the city to issue permits to new small businesses. Another was to create a public version of the Web site Walkscore.com, on which ordinary people rate the walkability of their cities.

His most ambitious plan was a proposal to move beyond public information into the deeper and possibly more profitable mine of social-media data. Every day, he said, there are 250,000 New York-centric posts on Twitter alone — some concerning trash complaints, others unsanitary restaurant conditions. "If Young & Rubicam can use tweets to sell you stuff," he hypothetically asked, "why can't the city use them to make you less sick?"

This makes civil libertarians uncomfortable, particularly at a time when the Police Department's chief Big Data project — its use of the Compstat system to guide stop-and-frisk — is being questioned by the courts. Mr. Flowers insists that he has put in place safeguards, like keystroke logs on his employees' computers, to ensure that information is not abused. Still, groups like the New York Civil Liberties Union say that they are watching public data mining with a guarded, if optimistic, eye.

"I think that the Bloomberg administration's attention to data has enormous potential for good," Donna Lieberman, the executive director of the union, said. "Obviously, it means that the city can make and tweak policies based on reality. But the potential for the selective use and release of data is one aspect that raises concern."

Another, at least for Mr. Flowers, is whether his geek squad will survive the end of Mr. Bloomberg's tech-friendly tenure. For now, he said, he is proceeding under the assumption that it will, adding that the best way to ensure its viability is to create an appetite among city agencies for the analytical work his group produces.

"We know that there will always be a Fire Department, a Finance Department, a Department of Buildings," Mr. Flowers said. "So hopefully by building a common data infrastructure that shares information in real time, it won't matter who sits in City Hall."

Working in his favor is the firm belief among information activists that Big Data's moment, especially in the management of cities, has powerfully and irreversibly arrived. This is a conviction based on certain technological advancements and a discernible shift in how the younger generation sees its relationship to government.

What used to be about passively receiving services and dictates is now about participation, said Jennifer Pahlka, the executive director of Code for America, a volunteer group of techies that helps city governments, including New York City's, write code for public projects.

"Young people, because of social media, have always felt they've had a voice," Ms. Pahlka said. "They're coming from the assumption that government is a hackable system — an operating system that can be optimized. It's in their DNA, and they just go and do it."

A version of this article appeared in print on March 24, 2013, on page MB1 of the New York edition with the headline: The Mayor's Geek Squad.