

CHAPTER 2

THE MAYOR'S OFFICE OF DATA ANALYTICS

Institutionalizing Analytical Excellence

CRAIG CAMPBELL, *Harvard University*

STEPHEN GOLDSMITH, *Harvard University*

Executive Summary

This chapter explores the inception and operating model of the New York City Mayor's Office of Data Analytics (MODA)—a young government unit whose story is fraught with identity crisis, having been variously termed New York City's "civic intelligence center" and City Hall's "skunk works," said to have been staffed by the "Mayor's Geek Squad" or, alternatively, a team of "data therapists."¹ These varying epithets reflect changes in leadership, staff makeup, and institutional design in the eight years since its inception. What persists throughout these changes, however, is that MODA is foundationally an in-house consultancy helping city agencies use data and analytical techniques to improve how they deliver on their missions. By combining a shared service model for public-sector analytics with a broad data stewardship mandate, MODA represents a structural innovation in the way municipal government uses data analytics to produce public value.

The term "data analytics" has been deployed in a variety of competing contexts; here, we define "operational data analytics" as the protocols for producing actionable knowledge at the intersection of two discourses. The first is *operations research*, a discipline with a long history in government that is primarily concerned with using data to conduct deep analysis into a specific logistical activity or area of service delivery (Pollock & Maltz 1994). The other is *data science*, the scientific inquiry into the technology and methods of data analysis (Donoho 2015).

MODA provides New York City government a unique set of capabilities through its three programmatic areas: providing advanced analytics services to other city agencies; facilitating data-sharing among city agencies; and administering the city's open data program. Its footprint in various projects has ranged from light-touch advisory roles to large-scale interventions in municipal operations, in areas from economic development (see chapter 4) to sustainability (see chapter 1) to tenant protections and emergency response.

Before it was formally coined "MODA," though, it was a small team pulling together data sets on mortgage fraud under the supervision of the mayor's criminal justice coordinator. A few key factors—executive support, early successes in inspections-use cases, and a message well suited to a swelling nationwide interest in data-driven government—led to MODA's official instantiation by then-mayor Michael Bloomberg in a 2013 executive order. In 2014, in a new administration and under the direction of a new chief analytics officer, MODA moved into the Mayor's Office of Operations and realigned its portfolio around a new set of administrative priorities (see figure 2.1).

Although the use of data analysis is certainly not a new phenomenon in municipal government, the organization of an analytics capacity in an agile internal-consulting model, combined with a data-stewardship responsibility for making information reusable across the city, represents a compelling approach to leveraging data analytics as a force multiplier in the production of public value.

Key Takeaways and Actionable Insights

- *Institutionalizing analytics as a service improves knowledge management across policy areas.* By building analytics into the city's institutional architecture, outside any single policy area or operating vertical, MODA creates capacity for leveraging analytical insights in one domain across other institutional functions and subject areas.
- *Internalizing data intelligence provides cost reductions across the enterprise.* Complicated data or analytics questions faced by governments are often outsourced to outside consultants. MODA provides this expertise at no cost to city agencies, while ensuring that all processes are transparent and well-documented.
- *Operational focus and rapid prototyping ensures high-value service delivery.* Because of the "actionable" posture of its analytics work, MODA prioritizes projects that have a high likelihood of resulting in changes that optimize city processes. Because of its penchant for piloting new ideas, MODA provides client agencies an experimental sandbox where failure and rapid iteration are built into the analytics process.
- *Analytical excellence within government drives civic engagement externally.* MODA is not only a center of excellence for the use of data within government; through its implementation of the de Blasio administration's *Open*

Data for All vision, MODA also serves as a model for connecting government-transparency initiatives with data-literacy efforts, empowering those outside of government to use city data to take action on their own behalf.

- *Public interest in open data drives better use of data internally.* Pairing data analytics with open data program administration in the same office, MODA channels public interest in open data to improve the quality of the data sets that are also used for analysis within the city.

Theoretical Background, Contextual Environment, and Problem Definition

How does a city decide to improve? In this case study, we will not attempt to answer this question wholesale, but instead explore it through the lens of one key part of the twenty-first-century government innovation agenda: data analytics.

Data analytics, as a field, can be defined as the rough art of making sense of a morass of information to improve how decisions are made. The case for data analytics in the private sector is clear: quantifying information and organizing an institutional aptitude to understand poses clear benefit to profit-seeking firms that define success in the empirical terms of a monetary bottom line. In the public sector, clear measures of success can be less straightforward. The question of what is valuable for a public administration to produce—and what costs are acceptable for realizing these goals—is more complex, clouded with ideology and often determined by more qualitative measures of success (Moore 1995). For government, the opportunity for data analytics is at once more promising and more difficult to effectively organize at the institutional level.

We will examine the Mayor's Office of Data Analytics as an example of an institutional innovation that sits at the center of several recent trends in public administration, including an increased understanding of the capacity of data science, a growing appetite for organizational innovation, and more sophisticated data management strategies.

The Analytics Opportunity for City Government

City agencies and offices collect and maintain data on a variety of activities, including operational information on licenses, service requests or complaints from the public, and capital planning and investments; administrative data on revenue and procurement, performance indicators, and survey data; and more recently, embedded-sensor, social-media, and crowdsourced data.

Taken as a whole, a municipality has an enormous amount of data at its disposal. It naturally follows that data analytics—the means to derive actionable

insight from otherwise latent information—has the potential to make a significant impact in public administration. At a more granular level, however, the way in which data analytics can be deployed is complicated by the institutional culture, bureaucratic design, real and imagined legal challenges, and legacy information-technology systems common to city governments.

When the organizational structures prevent information from being socialized with other units, it is commonly labeled a “data silo”—the information-storage equivalent to “service silos” that are seen as anathema to streamlined customer experience and internal collaboration in a variety of organizational contexts.² A frequent assumption is that by creating lines of access across these silos, the enterprise at large is suffused with the data maintained by each of its constituent pieces. As information is socialized, analytical insight and decision-making can, from any organizational locus, operate with a complete picture of the city.³

Silo-breaking alone, however, does not result in operational improvement. For its data to be used effectively, a city also requires a culture that identifies high-value use cases by challenging long-standing procedures and norms. This culture is constituted by technical capabilities enabling analysis of data, a formal capacity for identifying patterns in this information, and a delivery strategy that allows for pattern-finding to be acted upon.

One salient example in New York City is municipal emergency management. Preparing for and responding to major emergencies is a primary objective of several city government agencies. Emergencies activate city resources in often novel ways, forcing the creation of new data and operational workflows that may not have existed previously. They prompt unanticipated questions—for example, where are all commercial cooling towers in New York City?—that the city, prior to that moment, had never needed to systematically ask.

Accessing high-quality data to answer these new questions, in the face of rapidly evolving circumstances and conflicting or incomplete information, is precisely the kind of situation that demands the creation of new infrastructures, protocols, and organizational units—and a way to make sense of it all.

Enter the Mayor’s Office of Data Analytics.

Inception, Development, and Rollout: From Ideation to Implementation

From the Financial Crimes Task Force to MODA

Former New York City mayor Michael Bloomberg, a leader in technological innovation in both business and government, recognized the importance of harnessing data and analytics to improve how the city operates. Throughout his three-term tenure, a number of units across city government brewed their own

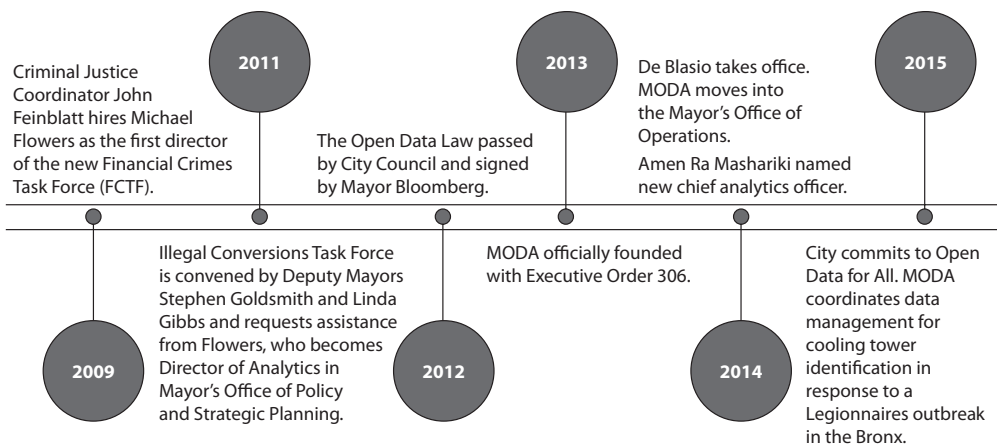


FIGURE 2.1 MODA timeline.

TABLE 2.1 Stakeholder Table

Stakeholder group	Goals	Role played	Opportunity for MODA
City leadership (mayor and deputy mayors)	Deliver on mayoral platform priorities more effectively and coordinate service delivery across institutional units using data and analytics	Supplied political will to found and fund an office of data analytics and a chief analytics officer	Allow leaders to “query” their city by using data to create better situational awareness; coordinate multiagency data efforts that fall outside any single domain
City agencies (e.g., Department of Buildings and City Commission on Human Rights)	Deliver value to New Yorkers based on their missions	Served as clients to MODA and supply data for analytics-use cases	Add measurable value to agencies’ operations by applying advanced data analytics to improve service delivery
Local data and tech community (e.g., BetaNYC, businesses using city data)	Create information products and services for New Yorkers	Provided feedback to improve city data through the Open Data initiative	Expand value proposition of city data by publishing it for public reuse through Open Data

data-centric solutions to government problems.⁴ One of them was the city's administration of the criminal justice system, managed by John Feinblatt, who served as the city's criminal justice coordinator and Bloomberg's chief liaison to district attorneys and the state criminal justice system from 2002 to 2013.

In 2009, Feinblatt hired Michael Flowers, a former Manhattan district attorney and Department of Justice military crimes investigator in Iraq, as the first-ever director of the newly formed Financial Crimes Task Force. The task force was created to coordinate the city's response to the 2008 financial crisis and started with no staff and no budget, just a mandate: Flowers was to identify and combine data from across city agencies to better understand the potentially criminal activity that contributed to the financial crisis, and then to hand this information off to the city's law enforcement agencies such as the district attorney's office.

Over the course of his first year on the job, Flowers built a small team of analysts and mapped out which agencies held relevant data, identifying in the process a use case prime for an analytical solution. Integrating and analyzing data—including property-deed transactions, various permits, and construction data—Flowers's team stitched together a picture of mortgage fraud that had contributed to irrational growth of the real-estate market. But while the model showed compelling evidence of illegal activities, the victims of fraud were big banks. Prosecuting on behalf of the major financial institutions was not, at the time, an urgent policy goal of the district attorney's office, and the mortgage-fraud project concluded without a single indictment.

This was a learning experience for the incipient analytics team. The process they had followed—identify a broad issue area, locate relevant data, specify a use case, and then match the evidence to an implementation pathway—needed to be reversed. Data analytics did not fix problems outright; for analytical solutions to be effective, the group needed to actively align incentives across a variety of stakeholders (table 2.1) to ensure that analytics deliverables would be acted on.

Meanwhile, other city leaders were pursuing their own data and analytics agendas. By mid-2010, the deputy mayor for operations was working to build a data analytics center and charged agency commissioners with surfacing analytical questions that, if answered, would change their operations. The deputy mayor for health and human services was developing a data and legal infrastructure to allow caseworkers to make decisions with a more holistic picture of a given constituent's encounters with different city agencies. Data-driven innovation was a shared goal of top officials, which brought needed visibility and support to Flowers' team.

Building inspections presented Flowers an opportunity ripe for analytics. The Department of Buildings (DOB), which is responsible for ensuring the structural stability of the city's built environment, oversees inspections. One aspect

of this mandate is identifying apartments that unscrupulous landlords have illegally renovated to accommodate more tenants than their city-sanctioned capacity. Known as “illegal conversions,” these properties were not only a potential quality-of-life problem; they also presented an outsize fire risk. After five people were killed in fires during the summer of 2011, Feinblatt, along with the deputy mayors for operations and health and human services, convened an interagency task force to address the issue (Moore et al. 2011).

Flowers seized the opportunity to provide quick and valuable data insights into strategies that might preempt future fires. He eventually became the director of analytics for the Mayor's Office of Policy and Strategic Planning, bringing with him the young analytics team he had developed. His team provided the analytics support to the Illegal Conversions Task Force, creating indicators for high-risk conversions from data on tax and bank foreclosure records and 3-1-1 complaints. The result was a model prioritizing which buildings to inspect. According to MODA's 2013 annual report, the model allowed DOB to identify up to 70% of the illegally converted buildings by sending joint teams of building and fire inspectors to just 30% of the complaints (Mayor's Office of Data Analytics 2013).

This strategy gained enough traction that the team, still operating as the Financial Crimes Task Force, was reconstituted as its own mayoral office. In April 2013, Bloomberg formalized the team as the city's “civic intelligence center” through Executive Order 306, which establishes the Mayor's Office of Data Analytics as the unit where data from across agencies is aggregated, analyzed, and turned into actionable insights. The order mandates MODA to address five areas of work: “Collaborative, Data-Driven Solutions, Citywide Data Platform, Oversight of Data Projects, Data Liaison Duties, Implementation of the Open Data Law” (City of New York 2013). It also created the chief analytics officer position, which Flowers stepped into.

The executive order, in short, chartered an office to create a formal institutional function for providing insight into city data sources, guiding the development of the infrastructure to share it, and enabling stakeholders across the city to leverage an advanced analytics capacity for rendering the data usable.

From Bloomberg to de Blasio

Late in 2013, Flowers stepped down from his post as chief analytics officer, leaving behind a legacy of an analytics office styled as a consulting unit—a long way from his original charge of investigating financial crimes.

Not long after, in January 2014, Bill de Blasio took over the mayoralty. De Blasio ran a campaign promising to bridge the gap in a “Tale of Two Cities,” contrasting the New York City of the privileged elite with the city experienced

by “everyday New Yorkers” (Chen 2013). The equity focus of the incoming administration was surrounded by a very different rhetoric of technology than the business-oriented pragmatism of Bloomberg. How this would apply to MODA was uncertain.

Among de Blasio’s first political appointees was Mindy Tarlow. A seasoned expert in both public administration and nonprofit leadership, Tarlow was named the director of the Mayor’s Office of Operations, a large team that provides project management support to interagency initiatives, researches and evaluates innovative programs, and measures city agencies on their commitments and key performance indicators. Among Tarlow’s primary charges from the mayor: “bring the Mayor’s Office of Operations back to the center of government” (Tarlow 2017).

Soon after Tarlow took the helm at Operations in January 2014, an auspicious turn of events brought some of the city’s most advanced data and analysis functions within her purview. The deputy mayor for health and human services agreed to move the Center for Economic Opportunity—the city’s in-house antipoverty think tank and a center of excellence for economic research and program evaluation—into Operations. Not long after, the research director of the Mayor’s Office of Criminal Justice similarly pivoted MODA into Operations.

Structurally, this represented an opportunity to improve data use citywide by leveraging Operations’ tactical toolkit. Through its performance-management role, Operations has a broad, lateral view of the city, providing a picture of how well the city is meeting its goals at any moment. Through a number of specialized “service” units—such as its project-management team and CEO—Operations engages tactically on a project basis as in-house consultants. Within this structure, MODA’s focus, given its small size and sharp mission, is to engage on those projects where advanced analytics can have the biggest impact.

With MODA in Operations in mid-2014, Tarlow began a search for a new chief analytics officer, looking for “someone who could combine an analytic mindset with an equity-orientation” (Tarlow 2017). She found a match in Amen Ra Mashariki, a Brooklyn native who brought to the role a doctorate in engineering, data-intensive roles in the telecommunication and biomedical fields, and a background in technology leadership in federal government as the chief technology officer of the federal Office of Personnel Management (McEnery 2014).

One of Mashariki’s first orders of business when he took the position in October 2014 was working with a small but astute set of holdovers from Flowers’s team to hire a set of four “aggressive analytical thinkers.”⁵ In light of a different staff makeup, a different set of administrative values, and a different organizational location in Operations, the team had to prove its value in a

new administrative climate. Mashariki's strategy: start small, collect "wins," and show how the cross-functional analytics team could thrive not only under the technology-driven management style of Michael Bloomberg, but in any environment (Mashariki 2017b). "We wouldn't be successful if we jumped into the scrum battles over the sexy tech projects," Mashariki said. "MODA could work with agencies on small things and get small wins."

In the summer of 2015, two pivotal events came to pass that would define MODA's future. The first was a deliberate act, in collaboration with the Department of Information Technology and Telecommunication (DoITT), to define a new direction for the Open Data program. The next was an unexpected public-health emergency that unearthed a deep need for a federated data-governance model and an analytics environment to support it.

Open Data for All

New York City began publishing administrative and operational data sets maintained by city agencies in 2009 with the inaugural NYC BigApps, a competition designed to spur community data science and app development with city data. NYC Open Data, from its inception, was seen as an opportunity to engage the technology community as a key resource for local entrepreneurialism. Two years later, the initiative was written into the administrative code when City Council passed Local Law 11 of 2012, also known as "the Open Data Law," which added an aggressive transparency provision to the overall mandate of open data. With the 2013 executive order establishing MODA, the role of the chief analytics officer also captured the responsibilities of the chief open platform officer—leader of the Open Data initiative (Campbell 2017).

In July 2015, three years and 1,300 data sets into the program, the de Blasio administration charted a new direction for the program. MODA and DoITT published *Open Data for All*, the strategic document that redoubled the program's focus on everyday New Yorkers and aligned the city's data-publishing program with the de Blasio equity agenda (City of New York 2015). Open data for the local technology elite was, by 2015, a well-understood use case. *Open Data for All* marked the start of the administration's effort to make open data a part of its community empowerment and digital literacy toolkit.

The *Open Data for All* strategy committed to building feedback into the data-disclosure process. It also set a course to make explicit the abstract idea of who was benefiting from public data. Starting with a Citywide Engagement Tour and segueing into a number of other user-research and community-engagement initiatives, MODA committed to doing more to understand who was actually using city data while simultaneously spurring greater use.

Cooling Towers and Legionnaires' Disease Response

The following month, several New Yorkers were killed during an outbreak of Legionnaires' disease, a respiratory infection caused by *Legionella* bacteria (Hu and Remnick 2015). When the Department of Health and Mental Hygiene (DOHMH) learned that these bacteria were incubating in cooling towers (HVAC systems on top of some buildings), the city suddenly encountered an urgent need to locate, count, and track inspections on all cooling towers. It was the kind of crisis the city had not encountered before (Chamberlain, Lehnert, and Berkelman 2017).

By early August, the initial outbreak in the South Bronx had been contained and the public-health emergency declared resolved by a quick response from the DOHMH. Although the crisis had subsided, it became a priority to mitigate the risk of *Legionella* bacteria in other cooling towers across the five boroughs. City leadership needed to locate all of the places where the bacteria might be incubating, test them for the bacteria, and then register that they had been cleaned. Time was in short supply: several New Yorkers' lives had already been lost.

A legislative solution was fast-tracked through City Council: Local Law 77 of 2015 required any building owner with a cooling tower to register and verify it had been inspected. With the statutory mandate in hand, the city deployed its "boots on the ground"—a combination of inspectors from the Department of Buildings and the Fire Department—and conducted proactive outreach to building owners through outbound calls and a site on nyc.gov to register buildings with cooling towers.

Given the breakneck pace at which the crisis was unfolding and the many moving parts of city government activated to respond, maintaining an accurate and complete picture of the status of inspections, in addition to routing new inspections efficiently, was imperative. What had started as an epidemiological problem became a data-management one. It was clear that infrastructure and protocol for handling a problem this complex and quickly evolving were not in place.

MODA, which was originally involved in providing information on landlord addresses, stepped in to help manage the data. Relevant agencies sent data to MODA, which acted as an analytics hub for the duration of the event. MODA analysts built a machine-learning model that reconciled counting methodologies from the various agencies involved and reported results back to City Hall within hours of receiving incoming data.

This emergency work, which necessitated immediate access to data, made it clear after the crisis subsided that MODA needed to help prepare the city for more coordinated data-sharing in response to emergency events. This included

developing an emergency-data model and accompanying data standards, as well as a protocol for mobilizing access to information in emergency situations—a mandate that came to define MODA's interagency data-sharing strategy in the years that followed.

Results and Lessons Learned

With its coordination of the cooling-tower-counting during the Bronx Legionnaires' outbreak, MODA demonstrated to the de Blasio leadership team the value of a data-analytics office outside any single agency or subject area. With *Open Data for All*, MODA gained visibility with the community of public users and created a message resonant with the mayor's equity agenda. These two formative actions created space to take the ad hoc nature of project-specific analytics work and formalize it into a coherent, consistent strategy for citywide data and analytics.

Data at the Speed of Thought

Data-management issues that surfaced during the Legionnaires' crisis prompted a new outlook on the role the Mayor's Office of Data Analytics should play in facilitating interagency data access. Tarlow (2017) observes:

Coming out of Legionnaires, we needed to have a way to build a foundation of all the information we have and a means to share it . . . so that when an emergency happens, we can focus just on the emergency itself. When it was cooling towers—or in the case of Zika, still or standing water—we should have all of the foundational information already and only be focusing on sharing relevant information we have and collecting any new information specific to the issue at hand.

According to Tarlow, a citywide crisis should never be an information crisis. Relevant information should be furnished at the “speed of thought,”⁶ meaning that as soon as a need for information is articulated, the infrastructure and protocol are already in place to furnish it immediately. Mashariki explains further that, while one-to-one agency data exchanges were common, the challenge lay in creating an integrated, citywide view.⁷

Early efforts in New York City to integrate data across a variety of sources were focused in the areas of health and human services and criminal justice. The value of data integration for these subjects is clear: matching records on a single person who accesses city services across a variety of departments and

information silos enables better coordination and more efficient service delivery. It also creates linked data ripe for analytics. “The same process we used with Health and Human Services, we could use with Emergency Management” (Tarlow 2017). Under Tarlow’s leadership, the Mayor’s Office of Operations led the effort to standardize policy and protocol, creating an omnibus policy for interagency data collaboration called the Citywide Data Integration Framework. In parallel, Mashariki set out to build the Citywide Intelligence Hub, a data-sharing platform that agencies across the city could use to share data and use collaborative applications.⁸

Among the city’s core data infrastructure assets is DataBridge, a central data warehouse managed by MODA, DoITT, and the Mayor’s Office of Operations that stores data sets shared between city agencies. The core innovation of DataBridge, when it was designed and implemented for citywide data retention and analytics, was that it enabled MODA and other key analytics teams to access data that was already being shared out of its source system. Its implementation and functional architecture resulted in a first-of-its-kind centralization effort.

Originally coined “DataBridge 2.0,” the hub strategy takes a federated, rather than centralized, approach to citywide data-sharing. The hub system does not store any data itself but provides a view directly to agency data systems, allowing users to search for and share data with each other through a user interface oriented around geographic tools. According to Mashariki (2017a), this approach better accommodates the fluidity of data assets: “If we got every single data set that exists in the City in [DataBridge], then by tomorrow there would be ten more that we are missing,” Mashariki said. A hub, alternatively, accommodates the fact that new data are created constantly, allowing data to be seamlessly drawn into other tools as soon as a need arises so that, as Tarlow put it, the data work happens in the background while officials “focus on just the emergency itself” (Tarlow 2017).

In collaboration with MODA, New York City Emergency Management (NYCEM) created an exercise modeled after emergency tabletops and live operational drills for the city’s data and analytics personnel, which leveraged the Intelligence Hub prototype to test the system’s functionality outside the context of an actual emergency. The first “data drill,” as the exercise came to be known, took place in July 2016 and simulated a large-scale power outage in Brooklyn. Agencies shared relevant data sets, which MODA analyzed to answer questions on how the boundaries of the blackout zone would be identified and how data could be organized to provide situational-awareness reports to city leadership. The drill showed that the hub was not only an effective data-sharing platform but also a good governance approach, convening relevant stakeholders at agencies while fostering a culture more prone to collaboration on data operations.

Beyond Data Publishing

Just as MODA helps make data accessible to city employees, it also supports efforts to provide that same data to a community of public users. MODA has encouraged and supported several implementations of *Open Data for All* as a force multiplier for existing agency initiatives. In 2015, for example, NYC Parks facilitated a participatory mapping campaign known as TreesCount! 2015, which mobilized 2,300 volunteers and more than sixty local organizations to count every curbside tree in the city. One year later, the NYC Parks analytics team worked with MODA and BetaNYC, a local civic hacker group, to host the TreesCount! Data Jam. Five of the community groups that collected data presented challenges to civic hackers, and the resulting projects were presented back to Parks leadership to consider taking on as operational solutions. The event also included a workshop for Parks volunteers and Open Data newcomers to develop skills to analyze TreesCount! data themselves—linking data stewardship with Parks' long history of tree stewardship. Parks used the *Open Data for All* framework to convene a coalition of stakeholders around a data set that had personal significance to a variety of existing affinity groups, including Parks staff, nonprofit volunteers, and civic tech partners.

In addition to empowering community users, Open Data serves a valuable function for city employees and citywide data management. The chief analytics officer is also the chief open platform officer, meaning the analytics principal also enjoys the organizational role of the city's director of open data. According to Mashariki, public interest and pressure foster a virtuous cycle that contributes to better data quality and governance in the city at large (Mashariki 2017c). MODA has used this position strategically, leveraging the community interest in the city's open data sets to advance an agenda for better data quality internally. For instance, an amendment to the city's Open Data Law required MODA to standardize street addresses and other geospatial attributes for locational data sets. MODA solicited public feedback on the standard and received a rich collection of expert opinions and actionable insights on ways to optimize geospatial data—not just for open data, as the law mandated, but also for internal use cases.

Advanced Analytics in Action

Sharing data—publicly or internally—is not inherently valuable. Analytics render data assets into the knowledge that, when acted on, creates public value. To this end, MODA structurally includes oversight roles over citywide data-sharing and open data (see figure 2.2).

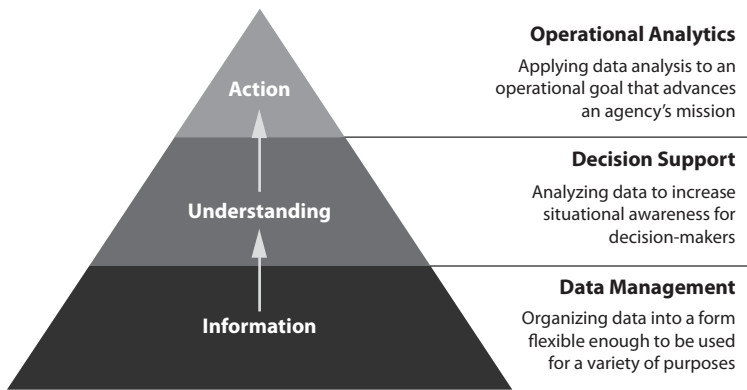


FIGURE 2.2 MODA project types.

MODA's analytics methodologies support the full span of a project, rather than isolating the "analytics" phase as a particular step in the process, and include not just programming, statistics, and other data-modeling methods but also business analysis, project management, stakeholder engagement, and policy analysis.

MODA's analytics process, which resembles the iterative techniques of agile project management, builds key benchmarks into the scoping, data management, analytics, and pilot stages of the project.⁹ These benchmarks pose simple questions—Is the problem solvable? Is there enough information? Is the model good enough?—which the MODA analysts and agency clients work together toward consensus.

According to Simon Rimmele, a MODA data scientist:

Each of these steps provides something of value. Just because a project doesn't go from one-to-five does not make it incomplete, and does not make it a failure. We encourage ourselves and others not to think that way. Projects may not begin at one and may not end at five, but provided we're completing each step in a way that's airtight, we've added value. Even if [a project] just reaches the scoping phase, if we help clients decide if a problem is solvable or unsolvable, we've helped the City. (Rimmele and Zirngibl 2017)

Many engagements start with an agency requesting data. In MODA's model, a data request represents a deeper need at the agency level and is an opportunity to investigate how analytics could add value. "We call it data therapy," says Rimmele. "The best way to start a project is to get everyone in the room. They state 'I want data set x.' Then we ask: is it just a data ask, or is there

a deeper problem? We go back and forth, asking probing questions to get to the real issues.”

In one example, MODA partnered with the City Commission on Human Rights (CCHR) to optimize where inspectors search for landlords who are illegally rejecting tenants on the basis of their source of income.¹⁰ CCHR first reached out to MODA for an exploratory conversation on data sets that could be applicable to the project. But working with MODA was an iterative process. According to Sapna Raj (2017), assistant commissioner for the Law Enforcement Bureau at CCHR, “We had many meetings where we said, ‘Oh this is not working, let’s tweak it a little bit.’ There was a lot of back and forth.”

The collaborative process paid off. Before MODA was involved, inspectors had to physically walk, block to block, building to building, to see where landlords were accepting new tenants. With the data, “there was a huge difference.” With the resulting analysis, the testers “didn’t have to do all that research to kind of figure out who owns it, who gets the tax bill. All those things became much more simplified than what they were doing before” (Raj 2017). These results were presented in straightforward maps that the eventual testers could understand without special training.

The analytics process is MODA’s framework, not only for delivering results, but also for familiarizing its partners and clients with analytical thinking. According to MODA data scientist Ryan Zirngibl, “a lot of what MODA does is not what the data was collected for—so you need smart people to give it the disclaimer” (Patel and Zirngibl 2016). The role of the MODA staff is “to challenge assumptions about the data as we massage data into being useful,” said Deena Patel (Patel and Zirngibl 2016), another data scientist at MODA. “It’s not just the outcome that matters—it’s a back-and-forth with the client where decisions are made and insights are gained from exploring the data.”

Challenges

“Should MODA be an *epicenter* of everything, or a *center of excellence*?” asks Mashariki (2017b). This question—whether MODA should administer the city’s collective analytics agenda or instead encourage innovation at the agency level—points to a core tension in MODA’s makeup and organizational placement, as illustrated in figure 2.3.

Data alone cannot drive operational innovation: a challenge facing any analytical organization is its ability to move data-driven insight into actual process change (Wiseman 2017). Clients interested in engaging a resource such as MODA may tend to be innovative thinkers who already see the value of data analytics. While this process, MODA analysts say, results in better uptake of analytical results, it may also result in fewer engagements with low-performing

Strengths	Weaknesses
Advanced analytics extract business value from otherwise latent information assets. Free in-house consultancy model with citywide focus allows for knowledge retention across subject areas, allowing the City as a whole to become more “intelligent” as more use cases are developed.	In-house consultancy may attract high performers, not areas where data and analytics could make the biggest impact. Institutional inertia in government creates barriers to moving projects from ideation and analysis to implementation.
Opportunities	Threats
Open Data opens co-production opportunities. Data issues and analytical insights delivered by public users can improve data upstream. MODA’s insight into data needs across City government allows the office to broker external data sources for internal use cases.	Risk-averse agencies may be reluctant to share data for citywide analytics projects. MODA’s placement in Mayor’s Office makes its existence and agenda subject to political lifecycle of the mayoralty. City data is accompanied by a complicated regulatory environment.

FIGURE 2.3 MODA SWOT analysis.

operations where analytical solutions may deliver outsized value—creating an open question as to whether MODA should ideate and pursue analytical questions before the relevant city agency generates one. Striking a balance between maintaining trust with clients and changing agency processes proves difficult, MODA’s data scientists say.

Culture Change: Conceptualizing Data as an Asset

MODA data scientists also described challenges guiding clients to see analytics as a problem solving tool, rather than a more facile view of “data analytics” as a particular technology or software product. Good technological tools can help, but they are always a means rather than an end. In the words of Zirngibl, “The agencies’ data is so useful that tech companies come to them all the time, trying to convince them that *their technology* will solve all of their problems. What they are really trying to sell you is the fact that *your* data and information is valuable—and that their tool can make sense of it” (Zirngibl 2016). This underscores the broader challenge in empowering agencies to manage their data as an analytical asset, rather than technology overhead.

Balancing Tactical and Strategic Priorities

In addition to its tactical work analytics projects, MODA plays a strategic role in advancing the analytics agenda of the city by allocating staff to advise on agency data strategy and build capacity for agencies to do their own analytics work. The balance between these tactical and strategic activities is a challenge, though, especially since indicators of success in “capacity building” significantly lag behind upfront investments. It is tempting to attribute the growing analytics culture throughout the city to MODA: several other city agencies have adopted MODA's analytics model, including the Fire Department, the Parks Department, and the Department of Buildings. In reality, however, the growth of analytics in city agencies parallels MODA's growth in a symbiotic way.

Conclusion

Several other cities in the United States and abroad have followed the model that MODA has pioneered. NOLytics, an analytics unit situated in the Office of Performance Management followed in MODA's footsteps in 2014. The idea for a London Office of Data Analytics (LODA) was directly inspired by MODA (Copeland 2015). And the Johns Hopkins Center for Government Excellence, as a consortium member of the Bloomberg Philanthropies' What Works Cities program, assists some of the network's one hundred midsize cities in developing their own dedicated data and analytics capacity in the style of MODA. Despite its lean team, MODA has become the standard-bearer in municipal analytics.

MODA's success owes to a well-honed repertoire of analytics tools that often avoid the cutting edge of data science techniques—the neural nets, cellular processes, artificial intelligence, and evolutionary algorithms that excite those unschooled in the specific policy instruments available for change—through no shortcoming of technical skill. Rather, MODA's techniques are deceptively simple, asking questions such as: What goes with what? How much? Where to look? What if? Data analytics, for MODA, is less about cultivating a black box of data science than making a growing body of civil servants comfortable with asking analytical questions and using analytical tools to drive policy and operational decisions. Analytics deliver sustained value only when it is embedded in the institutional fabric of the city—a lesson learned that has been reflected in the organizational strategy of MODA over the course of two mayoral administrations.

Mashariki says that MODA's long-term theory of change for the city is expanding analytics capabilities of city agencies through its analytics and data-sharing portfolio, ultimately building capacity to enable its own obsolescence. While that future is not yet in sight, MODA's institutional model

has demonstrated that raising the analytical tide floats all boats. From its beginnings as a niche innovation wedged between existing functions in city government, through its sustained delivery of value across different political environments, MODA's continued existence offers evidence that decision-making informed by advanced analytics is becoming standard fare in New York City government.

Notes

1. "Civic intelligence center": www.nyc.gov/analytics, accessed March 20, 2016; "Skunk works in City Hall": Michael Flowers, quoted in Stephen Goldsmith, "Digital Transformations: Wiring the Responsive City," Manhattan Institute Papers, June 2, 2014; "Mayor's Geek Squad": Alan Feuer, "The Mayor's Geek Squad," *New York Times*, March 23, 2013; "Data therapists": Amen Ra Mashariki, quoted in "Digital Technology and the City: Leveraging and Managing Technology to Improve Services," NYU Wagner Graduate School of Public Service, December 3, 2015.
2. On analogizing data silos to service silos, see Wilder-James 2016. Specifically, the rhetoric of breaking down data silos has been frequently discussed in the context of business analytics (Duncan, Selvage, and Judah 2016), public health informatics, and centralized government management (Lubell 2017).
3. Media historian Shannon Mattern (2017) writes that to boosters of data-driven cities this perspective "appeals because it frames the messiness of urban life as programmable and subject to rational order." She argues, however, that this approach risks overlooking "data's human, institutional, and technological creators, its curators, its preservers, its owners and brokers, its 'users,' its hackers and critics"—that is, the imperfections and political instrumentality of information.
4. For a full account of MODA's inception and early work, as well as the other data-driven initiatives happening in New York City government at the time, see "The Data-Smart City," chapter 6 of *The Responsive City: Engaging Communities Through Data-Smart Governance* (Goldsmith and Crawford 2014).
5. Mashariki (2016) added: "I hire people based less on how well they match a profile of technical skills, more on whether they are curious and aggressive analytical thinkers."
6. Mashariki attributes the idea of "emergency data at the speed of thought" to Henry Jackson, chief information officer at New York City Emergency Management. It references Bill Gates's *Business @ the Speed of Thought: Using a Digital Nervous System* (Grand Central Publishing 1999), which explains how managers should view technology not as overhead, but as a strategic asset.
7. Mashariki elaborates: "You have people who have been in government for 20, 30 years who've worked at four different agencies. And because they've worked at four different agencies they know the people to call at those agencies when they need stuff. So there is a mechanism that exists, [which is] 'Oh I know that guy at the agency, let me call him.' But there wasn't a framework or an infrastructure across the board" (quoted in Deleon 2017).
8. The core of the agile approach, which was proposed and refined in the software-development context, is adaptability to process that follows value, as opposed to protocol-driven process where value production is secondary to a plan and role. (Agile Alliance 2001 and Flora and Chande 2014).

9. For a more in-depth discussion of the interagency data hub approach, and its incarnations in different cities, see chapter 5 in *The New City O/S* (Goldsmith and Kleiman 2017).
10. For more information on the MODA-CCHR partnership, see McKenzie 2017.

References

- Agile Alliance. 2001. "Manifesto for Agile Software Development." Accessed October 1, 2017. <http://agilemanifesto.org/>.
- Campbell, Craig. 2017. "New York City Open Data: A Brief History." *Data-Smart City Solutions*, March 8, 2017. <http://datasmart.ash.harvard.edu/news/article/new-york-city-open-data-a-brief-history-991>.
- Chamberlain, Allison, Jonathan Lehnert, and Ruth Berkelman. 2017. "The 2015 New York City Legionnaires' Disease Outbreak: A Case Study on a History-Making Outbreak." *Journal of Public Health Management and Practice*, March 25. doi: 10.1097/PHH.0000000000000558.
- Chen, David. 2013. "De Blasio, Announcing Mayoral Bid, Pledges to Help People City Hall Forgot." *New York Times*, January 27. <http://www.nytimes.com/2013/01/28/nyregion/bill-de-blasio-kicks-off-campaign-for-mayor.html>.
- City of New York. 2013. "Executive Order 306: Establishing the Mayor's Office of Data Analytics." Signed by Mayor Michael Bloomberg on April 17, 2013.
- . 2015. "De Blasio Administration Releases Open Data for All, the City's New Open Data Plan." Press Release, July 15.
- Copeland, Eddie. 2015. "Big Data in the Big Apple: The Lessons London Can Learn from New York's Data-Driven Approach to Smart Cities." Capital City Foundation, June 2015. <http://eddiecopeland.me/wp-content/uploads/2015/11/Big-Data-in-the-Big-Apple-Report.pdf>.
- Deleon, Nicholas. 2017. "How Data Can Stop Your City from Burning to the Ground." *Motherboard*, May 10. https://motherboard.vice.com/en_us/article/data-can-stop-your-city-from-burning-to-the-ground-smart-cities-nyc.
- Donoho, David. 2015. "Fifty Years of Data Science." Proceedings of the Tukey Centennial Workshop, Princeton University, September 18, 2015.
- Duncan, Alan, Mei Yang Selvage, and Saul Judah. 2016. "How a Chief Data Officer Should Drive a Data Quality Program." Gartner Research, October 14. <https://www.gartner.com/document/3471560>.
- Flora, Harleen, and Swati Chande. 2014. "A Systematic Study on Agile Software Development Methodologies and Practices." *International Journal of Computer Science and Information Technologies* 5, no. 3: 3626–37.
- Goldsmith, Stephen, and Susan Crawford. 2014. *The Responsive City: Engaging Communities Through Data-Smart Governance*. Hoboken, NJ: Jossey-Bass.
- Goldsmith, Stephen, and Neil Kleiman. 2017. *A New City O/S: The Power of Open, Collaborative, and Distributed Governance*. Washington, DC: Brookings Institution Press.
- Hu, Winnie, and Noah Remnick. 2015. "A Belated Look at New York's Cooling Towers, Prime Suspect in Legionnaires' Outbreak." *New York Times*, August 4. https://www.nytimes.com/2015/08/05/nyregion/new-york-officials-move-to-regulate-towers-tied-to-legionnaires-disease-outbreak.html?_r=0.
- Lubell, Sam. 2017. "Virtual Singapore Looks Just Like Singapore IRL—but with More Data." *Wired*, February 21. <https://www.wired.com/2017/02/virtual-singapore-looks-just-like-singapore-irl-data/>.

- Mashariki, Amen Ra. 2016. Remarks at the Inaugural Convening of the Civic Analytics Network. Cambridge, MA, April 11–12, 2016.
- . 2017a. Interview by Craig Campbell. New York, March 7.
- . 2017b. Interview by Craig Campbell. New York, April 20.
- . 2017c. Testimony before City Council at Oversight Hearing on Open Data Examination and Verification. January 22, 2017.
- Mattern, Shannon. 2017. “A City Is Not a Computer.” *Places*, February. <https://placesjournal.org/article/a-city-is-not-a-computer/>.
- Mayor’s Office of Data Analytics. 2013. “NYC by the Numbers: Annual Report—2013.” December 2013. https://civicio.files.wordpress.com/2016/07/annual_report_2013.pdf.
- McEnery, Thornton. 2014. “De Blasio Names New Chief Analytics Officer.” *Crain’s New York Business*, October 22. <http://www.crainsnewyork.com/article/20141022/TECHNOLOGY/141029947/de-blasio-names-new-chief-analytics-officer>.
- McKenzie, Jessica. 2017. “How the NYC Commission on Human Rights Uses Open Data Against Discriminatory Landlords.” *Civicist*, July 13. <https://civichall.org/civicist/how-the-nyc-commission-on-human-rights-uses-open-data-against-discriminatory-landlords/>.
- Moore, Mark. 1995. *Creating Public Value: Strategic Management in Government*. Cambridge, MA: Harvard University Press.
- Moore, Tina, Brian Kates, Benjamin Lesser, and Greg Smith. 2011. “Recent Fatal Fires at Illegal Apartments Forcing City to Aggressively Attack Problem in Courts.” *New York Daily News*, May 22. <http://www.nydailynews.com/new-york/fatal-fires-illegal-apartments-forcing-city-aggressively-attack-problem-courts-article-1.143141>.
- Patel, Deena, and Ryan Zirngibl. 2016. Interview by Craig Campbell. New York, December 18.
- Pollock, Stephen, and Michael Maltz. 1994. “A Brief History of Operations Research in the Public Sector.” *Handbooks in Operations Research and Management Science* 6: 1–23.
- Raj, Sapna. 2017. Interview by Craig Campbell. New York, May 25.
- Rimmele, Simon, and Ryan Zirngibl. 2017. Phone interview by Craig Campbell. February 22.
- Tarlow, Mindy. 2017. Interview by Craig Campbell. New York, April 10.
- Wilder-James, Edd. 2016. “Breaking Down Data Silos.” *Harvard Business Review*, December 5.
- Wiseman, Jane. 2017. “Lessons from Leading CDOs: A Framework for Better Civic Analytics.” Ash Center for Democratic Governance and Innovation White Paper, January 2017. http://ash.harvard.edu/files/ash/files/leasons_from_leading_cdos.pdf.
- Zirngibl, Ryan. 2016. Interview by Craig Campbell. New York, September 22.