

MAY 2014

# Learning from New York City's open-data effort

Finding the technology needed to make data truly open is easy. Breaking down barriers to change and helping users understand the information takes work.

**Opening government data** demands much more than technology, says Mike Flowers, the former chief analytics officer of New York City. In this interview with McKinsey Global Institute partner Michael Chui, Flowers details the bridges that needed to be built, internally and externally, to make New York's open-data effort succeed. An edited transcript of Flowers' remarks follows.

## Cascading effect

The *Mayor's Management Report* is something that goes out every year. And it's been around for years, predating Bloomberg.<sup>1</sup> What Bloomberg did was make it extremely more robust. And what it is, really, is a series of KPIs<sup>2</sup> that the city reports on things it does. You could sit there and say, "Oh, we did X this or Y that," in terms of volume of inspections or widgets we delivered or whatever it is. But you have to be able to back up what you just said.

In order to do that, you need a back-end system that tells you how much you did. And that means you have to be tracking the system on a regular basis. It has this sort of cascading impact—the very imposition of these more robust KPIs. It doesn't tell you how to solve the problem, but it tells you where the problems are, or at least where you should start asking questions. Then we took the next step with analytics: figuring out how to solve those problems that we were now able to understand through KPIs. Now that might seem far afield from open data, but it's not. Open data, at least as it's practiced in New York City, is of a piece with our overall goal toward more effective city government.

<sup>1</sup> Michael Bloomberg was mayor of New York City from 2002 to 2013.

<sup>2</sup> Key performance indicators.

## Breaking down barriers

You've got technological, cultural, legal, and political hurdles involved. The technology part is the easy part. There are companies that provide that service, like Socrata. There are open-source systems, like CKAN. You can figure out a way.

If you don't even *have* that information in releasable form, then you've got a bigger problem than opening data. If somebody says, "Oh, we're going to start an open-data effort," and you walk into a warehouse full of boxes, then you have much bigger problems than getting an open-data website up and running. As for legal, I don't think it's a big problem if you're a good lawyer. You can figure something out. Most of this data is not statutorily protected. Perhaps it should be, and that's another conversation to have. But it's not right now.

It's really cultural and political—those are your two big barriers. And they're sort of tied together. Political stuff is really inconvenient. There are a lot of people who mouth transparency and say, "Yeah, we're going to release everything we've got." And then say, "Oh no, we're not. This is terrible. If I release our response time to this neighborhood versus this neighborhood, then we're just going to get a ton of heat from council person X or Y, that's going to get in the way of us doing our job."

And you know what? They're not wrong. That's true. Those calls will occur. And they're going to kick them off their ability to technocratically deliver that which they know how to deliver. But the answer is not to block off data. The answer is to release and explain it. Because forcing yourself to do that actually allows you to analyze internally just how good your processes are. I think that's a good thing. But, again, that's a political thing.

And then there's cultural. They're agencies that just aren't used to operating that way. At the end of the day, culturally speaking, you have to be made to understand that there's benefit for you if you put this data out. It's not just a pain. Culturally, this isn't just a pain in the butt for you. I understand it *is* a pain in the butt. But there's a net gain for you if you embrace this approach of transparency.

### Opening data isn't enough

Where I think a lot of people go off the rails with these kinds of projects is they dictate a set of solutions without any understanding of the on-the-ground realities of the agencies doing the work. By virtue of the approach that we took—by that I mean looking behind the data, looking to understand the processes that were behind the data—we were also in a great position to figure out a fix.

At that point, because I now knew how the Department of Buildings or the Department of Sanitation or the fire department or the water department went out and did what they did, and how they delivered their city widget, I could figure out the least disruptive place in the logistics chain to insert that insight, and then have it acted on.

An example here would be the Department of Buildings, when we changed their inspection system. That was actually a lot easier than it seemed because their inspections, the way they route them, they get a call from 311 for X or Y.<sup>3</sup> Then that 311 call is triaged by a 311 operator. That 311 operator puts it into the building information system, which is the very original name for their own IT system.

So we had to put that insight in between the 311 call and the entry into the building information system, because that's an old mainframe system. At that point, I just tagged every building receiving complaints with all of these other metrics, based on the algorithm that we had written, identifying higher or lower levels of risk. Automatically, it went out to the field as a prepopulated prioritization list. Didn't need to change a thing. So the people in the field didn't have to do anything differently in order for us to leverage what we now know.

### Help citizens understand data

When we release information and only a small number of users use that information, we've generated—as a government—an informational asymmetry. That's not cool. We shouldn't be doing that. We need to engage more broadly than we currently are, even if the people on the good end of that informational asymmetry are well-intentioned actors—whether they're civic activists or private citizens or academics or educators or whatever. It's still not right. That's a bad impact, in my view.

And then there are obvious problems, like a stalker. If I want to root around about person A, B, and C, it's infinitely easier now than it ever was. And then misunderstanding the information is the other problem that's big. It's just data. They're just LEGOs. It's not the building; they're bricks. It's not the building.

<sup>3</sup>New York City residents can dial the telephone number 311 to access nonemergency municipal services.

Somebody's got to understand the people and processes, the humanity and history behind those sales in the file we just put out on our open-data page if they're really going to leverage it the right way. Because it's extremely easy to misinterpret.

And I don't mean misinterpret in the sense of "Oh, it will make us look bad." I mean misinterpret—you're going to make a bad decision because you didn't understand the information you were relying on. I think we need to do a much better job of helping people understand that data, which means being much more transparent from a process-and-people standpoint and not just a data standpoint. Open data is a start. It's not the end. [□](#)

**Mike Flowers** is the former chief analytics officer for New York City. **Michael Chui** is a partner at the McKinsey Global Institute and is based in McKinsey's San Francisco office.

Copyright © 2014 McKinsey & Company. All rights reserved.