Testimony

John Kaehny, Executive Director, Reinvent Albany
Co-Chair, New York City Transparency Working Group

NYC Council Committee on Technology
Hearing on Open Data Audit Per Local Law 8 of 2016
January 24, 2017

Good Morning, Chairman Vacca and Members of the Technology Committee, I am John Kaehny, Executive Director of Reinvent Albany and Co-Chair of the NYC Transparency Working Group. Thank you for holding this oversight hearing and continuing to push the administration and its agencies to fulfill the promise and requirements of the Open Data Law from the beginning to the very end of your chairmanship.

This hearing is specifically about the Mayor's Office of Data Analytics, MODA, audit of the Department of Buildings, Department of Environment and Fire Department. The audit's larger purpose is to keep pushing open data forward by learning what is working and not working at some of the City's largest agencies.

Let me speak to that larger purpose for a moment. The City's open data program had a rough time making the transition from the Bloomberg to de Blasio administration and it took groups like ours making a ruckus with City Hall to get things back on track. We are again concerned about staffing and resources for open data.

Seven months ago, Rafael Carvalho, one of two senior open data managers at DOITT, left for another agency. He has not been replaced, and as far as we can tell there is no hiring process underway. Our take is that DOITT's inability or disinterest in replacing

this key open data position has slowed and undermined the Citywide open data effort. The person in this open data manager position is one of two people at DOITT who have the seniority to effectively work with, cajole, and massage agencies into getting their data published in a useful form. We ask that Council ask DOITT to fill this important, funded, position as soon as possible --- especially given the likelihood of citywide hiring freezes in the next budget.

The Audits

I have a number of general comments about the open data audits, but would note that we have had two business days to review the agency worksheets.

Broadly, we think the audits --- if done in the spirit they were intended --- can be extremely valuable tools for prodding specific agencies and for identifying systemic problems with open data implementation.

This said, MODA's audit is not nearly as useful as it could be. The audit report needs to include a concise summary of how well each audited agency is doing, what works well, what needs improvement and how each agency is going to meet its open data mandates. Without this, the audit is missing the forest for the trees and it is hard for even the informed public to know how healthy the open data forest is at the audited agencies. Put another way, without a lot of time and work, it's hard to tell which agencies are top open data performers and which are laggards.

This said, we think that this year's audit shows that the Department of Buildings is serious about open data and the Fire Department is not. We were impressed by DOB's detailed responses in the various audit worksheets, the number of datasets published and fact that vast majority of datasets have standard geocoding and data dictionaries.

In contrast, the Fire Department seems to have few datasets, lacks geocoding and data dictionaries and included sparse details about plans for improvement for each dataset, doesnt use the same data internally as it publishes and did not bother filling out a key worksheet which details which legally mandated reports are based on public data.

Like BetaNYC, we would also note that the Fire Department rejected a Department of Health request to publish more information about EMS dispatch data based on what we believe to be an incorrect assertion of privacy. (EMS dispatch data is available as open data in other cities, notably Seattle.) If our assessment of FDNY is unfair, we would like to know why.

Specific Questions from Audit

- What does MODA mean when it says that the term "dataset" is "loose." (page 4/17 Examination Report.)

- Related, which dataset does MODA believe agencies should be putting on the open data portal? Is this something that needs to be clarified in law?

- Did the audits find any examples of their own or other agencies using their agency data more effectively thanks to open data? If not, are they asserting that their data was easily accessible within their own agency and from other agencies?

- What did the audits tell MODA about how open data was making it easier for people within or at other agencies to access their data more easily?

- Do these agencies seem to think that open data helps them fulfill their mission?

- Do these agencies think open data has reduced their FOIL workload?

- Has MODA considered interviewing agency stakeholder groups when doing the audit to find out what they think?