Joint Testimony of

John Kaehny, Co-Chair NYC Transparency Working Group

Executive Director, Reinvent Albany AND

Gene Russianoff, Co-Chair NYC Transparency Working Group

Senior Staff Attorney, NYPIRG AND

Susan Lerner, Executive Director, Common Cause New York

Member NYC Transparency Working Group AND

Rachel Bloom, Director of Public Policy and Programs, Citizens Union

Member NYC Transparency Working Group

before the

New York City Council Committee on Technology

Hearing on Open Data on September 21, 2016

Good afternoon Chairman Vacca and Members of the Technology Committee, I am John Kaehny, co-chair of the New York City Transparency Working Group and Executive Director of Reinvent Albany. I am also speaking today on behalf of TWG co-chair and NYPIRG Senior Attorney, Gene Russianoff and Susan Lerner, Executive Director of Common Cause New York, and member of the TWG.

Thank you Chairman Vacca, the members of this Committee, and the Council for holding this extremely important oversight hearing, and for passing seven amendments to the Open Data Law. Your energetic support of Open Data have given a productive focus and momentum to the

City's efforts to make open data real and useful to the public and city agencies.

In 2012, New York City passed the world's first Open Data Law, and we are looked to as leaders by governments everywhere. New York City is serious about making open data work. New Yorkers make a huge investment in expensive city data and expect to get full value. It is working. We are beginning to see open data in everyday public life, with data sets like 311 and NYPD's and DOT's traffic crash data a regular feature of community board and city council meetings and news articles. Digital businesses are using Department of Building data to instantly inform building owners of violations, and transportation planners and the real estate industry are using huge TLC trip data sets to map neighborhood activity. Going forward, we expect open data to pay for itself many times over; city government will save tens of millions of dollars by making it much easier for city workers from one agency to find and use data from another, or from within their own agency.

In 2016, City Hall has made a vigorous effort to ensure that agencies will meet the requirements of the Open Data Law and address long-standing issues. Mindy Tarlow, the Director of the Mayor's Office of Operations, and DoITT Commissioner Anne Roest have helped get the City's Open Data Team more staff and more protected time to work on open data issues. This is particularly important because open data is only one of the Mayor's Office of Data Analytics four major responsibilities.

Notably, Tarlow and Roest have energetically reached out to open data stakeholders to address specific issues and to get advice on how to meet new open data mandates. Their open data team, led by Dr. Amen Ra Mashariki, the director of MODA, and Albert Webber at DOITT have begun to methodically address problems with geospatial standards, data dictionaries, and public feedback that were identified in previous oversight hearings. We appreciate their professionalism, pragmatism, and enthusiasm for open data.

Open Data Law Implementation 2016 Highlights and Issues

Before looking at what's next, here's what we see as highlights and low-lights of the administration's Open Data efforts in 2016.

2016 Highlights

1. DoITT gets an A+ for automating 100 data sets in 2016. That makes a total of 200 data sets out of 1,600 automatically updating on the Open Data Portal. This is great work, saves time and money, and improves the reliability of the data.

2. The administration gets an "A" grade for publishing important new data sets including the City Budget, City Record Online, Seven Major Felony crime data, and a huge TLC trip data set that was getting roughly 75 FOIL requests a year.

3. Thirty of eighty agencies reported on FOIL responses that included public data that is or should be on the Open Data portal. This is a good start.

4. DoITT and MODA staff are reading comments and requests on the Open Data Portal and responding.

5. The administration's Open Data Team published their annual update on-time and have done extensive analysis of open data usage.

2016 Outstanding issues

1. There is no clear process for fixing public complaints about data quality problems – and there are many.

2. City data remains non-standard and without data dictionaries.

3. Agency criteria for not publishing data remains very murky and there are still many high-demand data sets unpublished.

4. The Socrata open data platform has serious limits and should be re-assessed.

Moving Beyond Publishing New Data

So, what comes next and what can this Committee and City Council can do to keep open data moving forward? The big picture is that New York City is ahead of many other places, and Open Data here has entered a new phase. The first phase was all about publishing public data. Today, more than three-fourths of the City's public data has been published and the City is roughly on target to reach the 2018 publishing deadline in the Open Data Law. There is still plenty of work publishing data: only 200 or data sets out of 1600 are automatically updating. Plus, dozens of high-demand data sets like the NYPD's CompStat 2.0 crime data and DOT's street paving and milling schedules remain to be published as open data – but this foundational phase is mainly done.

Open Data 2.0: Data Quality, Standards, Fixing Errors, Public Feedback

New York City's next big open data challenge is to ensure Open Data is high-quality and easily usable. This means establishing standards and measurements and figuring out how the public can get data problems fixed. New Yorkers have been acting as free "data detectives"and have found dozens, possibly hundreds of significant errors in city data sets. Few are getting fixed.

Recommendations Technology Committee and Council

1. **Strengthen and clarify FOIL connection to Open Data**. Amend Local Law 7 of 2016 to require agencies to report to MODA and City Council the name of each data set requested and released under FOIL, and whether and when those data sets are scheduled for publication on the Open Data Portal.

2. **Consider mandating creation of public Data Issues tracker like the federal government's data.gov/issues**. This is a single site where data set errors are reported by public and are tracked with explanation of steps taken to get errors corrected.

3. **Continue holding oversight hearings; they are crucial**. And meet informally with the administration Open Data Team for regular updates. Council's involvement is essential.

Recommendations for De Blasio Administration

1. **Ensure the Open Data Team, especially MODA, have the time and resources they need to meet their open data responsibilities**. MODA staff have four strategic responsibilities, and open data is just one of them.

2. **Test a variety of ways for the public to report data quality and usability issues and get them fixed**. Try different approaches with different sizes and types of agencies, for example:

   a. Publish in Open Data Portal the phone number and email address of volunteer agency data set owners per the federal government's data.gov. (Note that the email and phone numbers of agency procurement officials are listed repeatedly in City Record Online and there have been no reports of problems.)

   b. Create public Data Issues tracker like federal government's data.gov/issues. This is a single site where data set errors reported by the public are tracked with explanation of steps taken to get errors corrected.

3. **Make Open Data implementation more transparent**.

   a. Clarify staff and budget for Open Data Team.

   b. Publish list of data sets scheduled to be automated by DOITT

   c. Publish known data issues in one place and report on status of efforts to fix them.

4. **Bolster the NYC Technical Standards Manual with additional guidelines for agencies publishing open data per New York State**. The state open data team has published very useful guidelines to

help agencies properly clean, format, and release data sets, with dozens of real-world examples and best practices.

Conclusion

Thank you Chairman Vacca and the Technology Committee for this opportunity to testify. On behalf of Reinvent Albany, NYPIRG, Common Cause/NY and the NYC Transparency Working Group, thank you again for holding this important oversight hearing. Your genuine interest and vigorous oversight are a fundamental reason that open data is working in New York City and continues to be a global model.