

DEPARTMENT OF INFORMATION TECHNOLOGY AND TELECOMMUNICATIONS
TESTIMONY BEFORE THE CITY COUNCIL COMMITTEE ON TECHNOLOGY
Oversight on The Open Data Law and its 2015/2016 Amendments
Wednesday, September 21, 2016

Thank you, Dr. Mashariki, and good afternoon Chairman Vacca and members of the City Council Committee on Technology. My name is Albert Webber, and I will be testifying today on behalf of the Department of Information Technology and Telecommunications (DoITT).

Overall Progress

As you know, this year's annual update to the NYC Open Data Plan was released on July 15, 2016. In this plan, DoITT, the Mayor's Office of Data Analytics (MODA), and City agencies collaborated to increase the quantity and quality of data available on the NYC Open Data Portal. Since last year's update, traffic on the Open Data Portal has increased, receiving over 5 million hits, with more than 2,000 user-created views being generated. To date, nearly 1,600 data sets have been made available, 116 of which were released in the past year (as per the Open Data Plan). These newly released datasets include Rodent Inspections, the City Record Online, and the City Council's own Constituent Services. The infrastructure has been built for the automation of over 100 new data sets, bringing the total number of automatically refreshed data sets to over 200. Additionally, we have identified and published more than 40 unscheduled data sets, which demonstrates the commitment of agencies to make data available as quickly as possible.

To continue building on the past year's accomplishments, DoITT is actively seeking to expand its open data team. We are currently interviewing candidates for three of four new positions that will be fully dedicated to the Open Data Program. With this additional staff, we plan to improve, among other things, our communication with Open Data users, our technical infrastructure, and the quality of the City's data.

The package of laws the City Council passed late last year and earlier this year have been an integral part of Open Data's advancement. The following portion of testimony will outline our progress with each law.

Local Law 106 of 2015: Data Retention

This law requires DoITT to preserve record level data, and to establish guidelines for the division of large data sets into smaller ones. The new guidelines will facilitate archiving of historical data as datasets get updated, allow users to follow data trends over time, and to improve the overall user experience. We are ahead of schedule for complying with this legislation. We have already begun working with agencies so data is preserved through maintaining active and inactive lists for some of our more popular datasets, appending instead of replacing existing records, and dividing some of the larger data sets into smaller, more manageable data sets. The Open Data team has already drafted the standards for data retention, and will open them up for public feedback later this fall.

Local Law 107 of 2015: Data Dictionaries

This law requires DoITT to include a plain language data dictionary with every data set on the Open Data Portal. Data dictionaries make data more understandable for our users, and fits into the vision of Open Data for All. To help ensure that data dictionaries are of use to the public, we plan to collaborate with users to create data dictionary templates, and to open our data dictionary standards for public comment later this fall. Since the law was enacted, we have already added approximately 70 data dictionaries, and will work towards including data dictionaries for all datasets on the Open Data Portal by the end of 2017.

Local Law 108 of 2015: Geospatial/Address Standards

This law requires the creation of a standardized geospatial address field layout for all Open Data datasets containing address information. This standard will not only result in more robust data being added to the portal, but will also allow users to easily perform spatial analysis across multiple data sets, and build applications that leverage this new information. As previously mentioned, Mayor de Blasio designated Dr. Amen Ra Mashariki, Chief Analytics Officer, to convene a working group which will provide recommendations for creating uniform geospatial standards to both DoITT and the City Council. To develop these standards, we held several meetings with Open Data Coordinators (ODCs), agency stakeholders, the Department of City Planning's Geosupport team, and Open Data Advocates. We then gathered this feedback, drafted standards, and opened it up for public comment earlier this month. We look forward to finalizing these standards, and providing the technical guidance to agencies to adhere to this legislation.

Local Law 109 of 2015: Timely Responses to Public Requests

This law requires DoITT to provide initial responses to requests for public data set within two weeks of receipt, and the agency to which DoITT referred a request to make a final determination within two months of receipt.

Over the past year the Open Data team has worked closely with agencies to respond to data requests dating as far back as 2011. We found that many of these requests were for data sets that already existed, reinforcing the notion that users have a true knowledge of the cities unique data landscape, and that agencies have been proactive about the release of data.

Improving feedback will continue to be a priority of the Open Data team moving forward. In the coming months we will be improving and consolidating our communication channels to make sure there is a clear line of communication between us and our users. We invite the public to continue requesting data sets, ask us questions, and comment on our proposed data standards.

Local Law 110 of 2015: Timely Dataset Updates

This law requires data sets updated on agency websites to also be updated on the Open Data Portal, and fits with our strategy to provide regularly refreshed and accurate data. To ensure timely updates, our

team will work with agencies to ensure there is a schedule in place to maintain up-to-date data on the Portal.

Regarding manual updates, our new staff will play a major role in ensuring that data is refreshed in a timely manner. We will also work to identify data on agency websites that will require manual updates. There are roughly 1,400 datasets that are not automatically updated and we will work to ensure they are all addressed.

Additionally, over the last year our technical team has made huge strides in making timely dataset updates. For the first time, we began the process of automatically updating geospatial datasets, a process that will save us hours of work every year. Additionally, partnering with the Department of Sanitation, our team will be releasing the raw data sourced from the PlowNYC application, which has been in high demand for many years. The data will be made available during snow events, with plow locations being updated several times per hour, and older records being archived. This is the first time we will be able to deliver automated data more frequent than daily, and we hope to apply these technology enhancements to other high-demand datasets.

Local Law 7 of 2016: FOIL

This law requires agencies to review responses to Freedom of Information Law (FOIL) requests that included release of data to determine if such responses include public data sets that have not yet been included on the Open Data Portal. The results of this review were included in the last update to the Open Data Plan, and will be included in updates going forward.

To facilitate agency compliance, DoITT's Open Data and legal staff provided guidance to agency ODCs and Counsel's office that clearly laid out what agencies needed to consider when reviewing their FOIL requests. This legislation was extremely helpful in expanding Open Data discussions within agencies and highlighting the association between data that is requested through FOIL and Open Data, both of which are aimed at promoting transparency. Reporting on these metrics included input from ODCs, FOIL officers, General Counsels, and Communications officers, and we anticipate it will continue to help agencies prioritize data going forward.

Local Law 8 of 2016: Examination and Verification

Last, but certainly not least, Local Law 8 of 2016 required the Mayor to appoint an office or agency to conduct a series of examinations and verifications of individual agencies' compliance with the Open Data Law. MODA is that appointed office, and began the process of collaborating with the Department of Investigation and its Commissioner on an examination and verification plan.

All agencies that must comply this year (the Department of Housing Preservation and Development, the Department of Correction, and the Department of Sanitation) have been briefed on the examination and verification schedule, and MODA will deliver surveys to each agency by the end of September.

Additionally, suggestions from the public will be collected throughout the fall.

Thank you for the opportunity to testify this afternoon. These laws have succeeded in making the Open Data Law stronger, and we have been pleased to update you on our progress. Open Data remains a priority of this administration, and we thank our partners in the City Council, and in the civic tech community for their continued advocacy.

This concludes our prepared testimony and we look forward to answering your questions.