



Testimony of Amen Ra Mashariki
Chief Analytics Officer and Director of the Mayor's Office of Data Analytics
Before the New York City Council Committee on Technology concerning

**Oversight – Open Data Plan 2015 and
7 Pieces of Legislation**

1:00pm, Thursday, October 1, 2015

14th Floor Committee Room, 250 Broadway New York, NY

Good afternoon. My name is Dr. Amen Ra Mashariki. I am the Chief Analytics Officer and Director of the Mayor's Office of Data Analytics (MODA), as well as the Chief Open Platform Officer for the City of New York. Here with me is Albert Webber of the Department of Information Technology and Telecommunications' (DoITT) Office of Advanced Technologies. The Mayor's Office of Data Analytics is housed within the Mayor's Office of Operations, and works in partnership with Anne Roest, the City's Chief Information Officer and DoITT Commissioner, to fulfill the requirements of the City's open data law, which was enacted as Local Law 11 in 2012. The Administration fully supports that law and works tirelessly to ensure it is fully implemented.

The promise of Open Data is transparent, free, and accessible data. Data is more than just numbers. It's information that can create new opportunities and level the playing field for New Yorkers. It illuminates issues to invite changes in frameworks, and leads to insights that turn impenetrable issues into solvable problems. It's the invitation for more voices to join critical conversations.

I am here today to testify on the progress of the implementation of the open data law via our reinvigorated interactions with agencies and our open data vision and strategy, as well as on the proposed pieces of legislation by the City Council designed to enhance the open data law. Over the past year, MODA and DoITT have been working energetically to engage agencies in continuing to open data sets to the public. To date we have released 1,386 data sets this year, up from 1,268 last year. This includes a number of notable new sets like trip data from the Taxi and Limousine Commission, data from City Record Online, and "Pre-K for All" data. Along these lines, we have and will continue to work diligently with agencies to assist them in releasing their data.

New York City remains the gold standard in producing open data, both nationally and internationally. Our ability to work with agencies to open their data to the public puts the City far and away ahead of all cities with respects to the number of public data sets available on a portal. Our open data efforts have also earned New York City a top spot on the Sunlight Foundation's national rankings of Open Data cities. While we believe this, in and of itself, is a strong accomplishment, the open data law is not just about releasing open data sets. As stated in the

Open Data law’s declaration of legislative findings and intent, open data serves to “streamline intra-governmental and inter-governmental communication and interoperability, permit the public to assist in identifying efficient solutions for government, promote innovative strategies for social progress, and create economic opportunities.” MODA and DoITT are partnering to realize a vision and strategy that will allow us to achieve those stated outcomes of the law.

Strategy and Vision: Open Data for All

“Open Data for All.” With this initiative we are saying that open data is an invitation for anyone, anywhere, at any time to engage with New York City government. The Open Data Plan, which was released on July 16, 2015, establishes the City’s vision to deliver on the promise of the Open Data for All initiative to the benefit of all New Yorkers. This is our vision of what we can do and what we aim to achieve with open data. The ultimate success of the Open Data for All initiative will not only be measured in the number of data sets that are published on the Open Data Portal, but in the number of New Yorkers who use open data in their daily lives. And that’s not just the tech-savvy New Yorkers – it’s all New Yorkers, in all five boroughs.

To deliver on this promise, we will:

- Continue to release new data sets and identify opportunities to provide data for all types of users;
- Focus on delivering quality data, enhancing the usability of the data portal, and expanding access to open data through use; and
- Continue to improve and enhance the technical infrastructure that provides the foundation for the entire initiative, and increase responsiveness to and flexibility for user needs.

With these goals in mind we are driving aggressively towards building an open data ecosystem that supports expanded access to open data and better quality of data, as well as enhanced portal usability.

Expand Access: Among the primary goals of the Open Data for All initiative is to meet users where they are. This year MODA and DoITT have begun a citywide open data engagement tour. This has included visits to CUNY campuses, visits to Borough Board meetings organized by the various Borough Presidents, online webinars, Open Data tutorials, and other community-specific speaking engagements. We’ll be gathering feedback from the citywide engagement tour to determine how to best meet the needs of all New Yorkers.

Deliver Quality: We also are ensuring data quality through the development of applications such as the Business Atlas. The Business Atlas is a portal that was created through a partnership between the Mayor’s Office of Data Analytics and the Department of Small Business Services. It uses open data and also partners with a small startup here in the city called Placemeter to provide market research for local entrepreneurs.

Enhance Usability: Earlier this summer we partnered with our open data platform provider to launch a new feature called Data Lens. Data Lens utilizes artificial intelligence technology in order to display data to everyday New Yorkers in a way that makes sense to them and is visually intuitive.

The Proposed Bills

The proposed bills amending the open data law are all focused on our shared goal of improving the quality, usability, and accessibility of open data.

MODA and DoITT have long believed that the Open Data Portal should be the primary location where agency data sets should be available, machine readable, and accessible. We have engaged agencies with this in mind. We understand that in order to do this, many costly mechanisms and processes must be in place. We firmly believe in working towards the goal set forth in Intro 915 to accomplish this. MODA and DoITT look forward to discussing with the Council how to best put a comprehensive plan in place to achieve this public data set updating capability.

We believe that in order for us to truly achieve a ubiquitous open data community, MODA and DoITT must be responsive to users as well as provide quality customer service to users. Timely responses to public requests on the Open Data Portal are paramount to us achieving this level of customer service. Intro 914 is very much in line with that goal and would enhance our responsiveness to such requests, and MODA fully supports the intent of this legislation.

MODA already instructs agencies to look at data that they have released in response to a FOIL request to decide whether the released data should be considered a public data set to be included on the Open Data Portal. Commissioner Toole of the Department of Records and Information Systems (DORIS) is currently building an Open FOIL system. We have discussed building a data exchange with this system that feeds directly into the Open Data Portal so that we can get automatic feeds of FOILED data. Such an approach would accomplish the goals set forth in Intro 908.

Enhancing the usability of the Open Data Portal is a key objective for MODA and DoITT. Ensuring that data dictionaries are available for each data set within the portal, as envisioned in Intro 898, will enhance the user's ability to understand and use the data. This is central to what we are trying to accomplish when we say Open Data for All. We support the intent of this bill and look forward to working with DoITT and the Council to identify the most appropriate way to implement this capability.

We believe that managing data standards throughout the portal is extremely important. We believe both high quality data dictionaries and the use of open data standards throughout the portal are achievable with agency and open data coordinator (ODC) engagement. Intro 900 is a great way to begin the conversation around citywide standardization of geospatial data, but would ultimately require processes that will require the expertise of the full tech leadership of the city, intimate agency engagement, and the investment of significant resources.

To ensure usability and access to all open data we are going to add language to the technology standards manual around archiving and maintaining old records instead of purging them after a

certain time period. Intro 890 similarly seeks to accomplish a strategy for the preservation of historical records. We must maintain a balance between keeping a clean, easy-to-use portal, preserving important historical data, and the cost of storing significant amounts of data that may have outlived their usefulness.

With respect to Intro 916, we have legal and practical concerns with the audit proposed by the bill. In particular, it appears that the bill would authorize the Commissioner of the Department of Investigation to inspect all of an agency's data, records, and information—including sensitive, confidential, and privileged information—in order to determine whether the information comprises a “public data set.” Although we support the principle of developing mechanisms to ensure compliance with the open data law, we do not think the proposed audit is the best way to achieve that goal. We are open to discussing alternative means for advancing agency compliance and accountability.

I would like to thank the Council Technology Committee for providing the opportunity to testify today and for your continued support of open data. At this point I'd like to turn it over to Albert Webber from the Department of Information Technology and Telecommunications to walk you through some of the finer details of implementing the open data law and the new enhancements we have released in response to feedback from the community. After Mr. Webber's testimony we will be happy to answer any questions you may have.