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.

INTRODUCTORY REMARKS

Good afternoon. My name is Dr. Amen Ra Mashariki. I'm the Chief Analytics Officer and Chief Open Platform Officer for the City of New York, as well as the Director of the Mayor's Office of Data Analytics, or MODA. I'm joined by Albert Webber of the Department of Information Technology and Telecommunications. We are here today to speak about the way MODA and DoITT have worked together over the last year to fulfill the requirements of the City's Open Data Law and its recent updates.

When city leaders passed Local Law 11 in 2012, New York City set a high bar for effective and transparent government everywhere. Not only would we open our data, we would open it *all*. Last July, we committed to Open Data *for* all, a more inclusive vision in which *all* New Yorkers, not just the tech-savvy, can find value in *all* the data. We believe that every New Yorker can benefit from Open Data, and Open Data can benefit from every New Yorker.

Soon after we published *Open Data for All,* Council passed and Mayor de Blasio signed a package of bills bolstering the original Open Data Law with requirements to make it easier for New Yorkers to access, understand, and derive value from City data online. These new laws required us to be more responsive to public feedback, create better technical standards on geospatial data and data retention, and strengthen our means of ensuring agency compliance. Together, they help anchor the Administration's commitment to transparency and equitable uses of technology around Open Data.

MODA is the business owner of the initiative. As a hub of analytics in the City, MODA advocates for the use of Open Data in citywide data analytics. DoITT is the technical owner, ensuring that technological capabilities are constantly improving to better meet user needs. Agencies are the data owners. Our primary points of contact with agencies are their Open Data Coordinators, with whom we work closely to ensure that agencies comply with the law.

These three entities, along with our friends in the civic tech community, help to form the City's Open Data team. Together, we continue to improve the biggest and best municipal Open Data initiative in the world by focusing on one number: 8.4 million, the total number of New Yorkers.

Over the last year, the new statutory mandates have catalyzed an alignment of our vision of Open Data for All with our technical tools and work processes. In order to implement these laws while growing the usability, reach, and impact of Open Data, we have drilled down on the following three areas:
- User research
- Feedback mechanisms
- Technical standards

USER RESEARCH

The first step in opening data should be focusing on demand. We support creative analytic thinking to enable our users to answer the questions most important to them, but we also hope to spur demand in communities less familiar to Open Data. *Open Data for All* means Open Data for app developers, agency analysts, and mapping gurus, but also Open Data for Emma Yang, the twelve-year old New Yorker who published her analysis of NYPD data from the Open Data Portal as the youngest-ever recipient of Wolfram Research's Mentorship Program.

To this end, MODA has partnered with the academic community to examine where Open Data is reaching New Yorkers – and where it's not. A capstone at the NYU Center for Urban Science + Progress measured "data poverty" across the City. Data poverty refers to a condition in which communities or people lack access to, use of, or representation within data that is nevertheless used to inform decisions that may affect them. This analysis is among the first of its kind and will be released publicly later this fall.

We also worked with the Columbia University School of International & Public Affairs to assess existing and future opportunities for Open Data to advance the goals of the City's Community-Based Organizations (CBOs). Although municipal data is in high demand, CBOs pointed to several pain points in accessing and operationalizing Open Data.

These results are informing conversations on ways to improve user experience on the Portal. Local Law 107 of 2015 is helping ensure that data is more understandable to non-technical users by requiring that every dataset include a data dictionary defining the attributes (and clarifying the usability) of its contents.

FEEDBACK MECHANISMS

The Open Data Portal is not a static product: it is a platform that should be used as a tool for public feedback. We are dedicated to soliciting user input at every step of the Open Data lifecycle to improve overall data quality and increase data use. We want those who give feedback to understand what their contributions are being used for, what opportunities for change are available, and when they can expect to see the changes take effect. Substantive engagement is more than an obligation: it must be default for Open Data.

This summer, we reviewed every comment that has ever been submitted to the Open Data portal and developed a typology for the types of inquiries that come in. This has informed how we are categorizing future feedback in order to optimize the way we deliver customer service.

Over the next year, we will be implementing a new technological solution to intake feedback with a centralized mechanism that allows for quicker responses and better tracking on our interactions with users. This will make it easier for us to meet the requirements of Local Law 109 of 2015, which

mandates timely responses to public requests. A better feedback mechanism gives us a better way to comply with the law.

This spring, as per Local Law 8 of 2016 on examination and verification, we worked with the Department of Investigation to develop a plan for examining the Housing and Preservation Department, the Department of Sanitation, and the Department of Correction for information assets that may contain public data.

Before this year, we would work with agencies to craft publishing plans for their data. Now, the examinations and verification law, the public requests law, and the law requiring coordination between Open Data Coordinators and FOIL Officers have enabled new means of locating data covered by the original Open Data Law. They also help enact a broader culture shift: as Open Data becomes the norm for city data, it makes agencies more aware of the data they have and the data they produce – and spurs better upkeep and disclosure of information.

TECHNICAL STANDARDS

Many of the new pieces of legislation require updates to the technical standards manual, the document of record that outlines all policies and protocols for Open Data. As we evaluate each standard, we have three priorities: meeting the letter of the law, meeting the spirit of the law, and maintaining usability to the maximum number of users.

To do this effectively, public participation is vital. Take for example Local Law 108 of 2015, which mandates the formation of a working group to create standards for address and geospatial information on the Open Data Portal. Over the last several months, I convened a group of Open Data Coordinators to draft recommendations for geospatial attributes, column headers, and geocoding processes for datasets. The recommended geospatial standards reflect the data fields most frequently captured by City agencies, information that is in highest demand from public users, and attributes that will have the biggest impact on citywide operations once they are standardized.

In the spirit of Open Data for All, we invited the public to join the conversation. We received constructive feedback and helpful suggestions from members of the tech startups, national non-profits, research universities, and others. For example, one submission from the American Society for the Prevention of Cruelty to Animals (ASPCA) said: "These new standards are a great step in the right direction and will allow nonprofits like ours to avoid costly geocoding of addresses."

These comments are helping us refine our recommendations to the standard, which will ultimately be included an updated version of the Technical Standards Manual later this year.

OPEN DATA IN ACTION

I'd like to close by sharing a standout account of the enormous potential of Open Data.

This spring, Ben Wellington, a local data scientist, studied parking violation data from the Department of Finance and found that some of the locations where parked cars were ticketed were in fact legal parking zones. When he brought his analysis to the attention of the City, NYPD analysts confirmed his results – also using Open Data – and took steps to ensure that ticketing officers would avoid the oversight in the future.

Wellington, who runs the popular analytics blog *I Quant NY*, wrote the following when he saw PD's response:

> "I was speechless. THIS is what the future of government could look like one day. THIS is what Open Data is all about... Imagine a city where all agencies embrace this sort of analysis instead of deflect and hide from it. Democracies provide pathways for government to learn from their citizens. Open data makes those pathways so much more powerful. In this case, the NYPD acknowledged the mistake, is retraining its officers and is putting in monitoring to limit this type of erroneous ticketing from happening in the future. In doing so, they have shown that they are ready and willing to work with the people of the city. And what better gift can we get from Open Data than that."

CONCLUDING REMARKS

Open Data lies at the nexus of digital services, public transparency, and cutting-edge civic analytics – and we are constantly aiming higher. We will continue to test our assumptions and try new ideas, collect information on our efforts, and analyze our performance for points of improvement. We have committed to engaging the public throughout the process to keep us honest and on track as we continue to learn and improve.

I would like to thank the Council Technology Committee for the opportunity to testify today and for its continued support of Open Data. At this point, I'd like to turn it over to Albert Webber, who will walk you through our progress on each piece of legislation individually, in order.

Testimony of Gale A. Brewer, Manhattan Borough President
Regarding New York City's Open Data Portal
New York City Council Committee on Technology
September 21, 2016

My name is Gale Brewer, and I am the Manhattan Borough President. I want to thank Chairman Vacca and the members of the Committee on Technology for holding this hearing. As you may know, I was the primary sponsor of Local Law 11 of 2012, the Open Data Law, as a member of the City Council. As Borough President, I am working closely with the Mayor to ensure its adequate implementation.

Implementing the Open Data Law continues to be a major undertaking, and I would like to thank Department of Information Technology and Telecommunications (DoITT) Commissioner Anne M. Roest and Chief Analytics Officer Dr. Amen Ra Mashariki for their efforts. The excellent work of DoITT, the Mayor's Office of Data Analytics (MODA), and the Mayor's Office of Technology and Innovation make me confident that New York City will continue to lead the nation in municipal data initiatives.

I am proud of our city's influential position in open data and the thriving civic hacker community that is leveraging the many opportunities created by transparency. I want to discuss a few examples of how open data is revolutionizing the ways in which citizens and city government interact.

Earlier this year, civic hacker Ben Wellington noticed that he kept getting ticketed for parking in front of sidewalk pedestrian ramps, even though a 2008 law sponsored by Councilman Gentile made such parking legal. Using NYC's Open Data portal, he looked at the city parking spots where such violations commonly occurred and found that the NYPD had issued over $1.7 million in tickets for parking in these legal spots. When he contacted our office with his findings, we connected him to the appropriate office at the NYPD. Within a few weeks, the NYPD reported that it had sent a training message to all officers clarifying the 2008 rule change and communicated this to precinct commanders with the highest number of summonses. The NYPD immediately worked to remedy this problem and is now monitoring these ticketing patterns. This example shows the power of open data to create new pathways for government to learn from citizens and for communities to realize they can make positive change happen.

Open data also strengthens New York City's economy. The Mind My Business app by Vizalytics Technology makes small businesses more resilient by notifying the business owner whenever someone files a 311 complaint on noise, trash, or cleanliness so that the owner can be proactive about improving operations and avoiding a fine. This innovation enables citizens to use 311 to listen to their patrons and run better businesses. The more the city communicates through its data, the more everyone stands to gain. Accordingly, when datasets are updated with greater frequency, their data are more actionable. The Open Data Law is an engine for economic development, and we've only beginning to see its full potential.

Open data is also a learning catalyst for New Yorkers in middle and high school. It can deepen students' understanding of the neighborhoods where they live, go to school, and hang out with their friends. Open data will be fueling the Data Science curriculum that my office is developing through our Data Science Education Task Force. This curriculum will teach students skills they can apply in many fields, and more datasets across more agencies will expand their universe of study. We expect to pilot this new course in two high schools in the 2017-2018 school year.

From watching and supporting the development of Manhattan's tech industry in communities like Harlem and Flatiron, I know that visionaries in this world learn by jumping in and getting their hands dirty. That's why for the third year in a row I am welcoming CUNY Service Fellows to connect the achievements of civic tech to Manhattan communities. These fellows are undergraduates from CUNY schools citywide and are learning about governance, community participation, data science, and of course open data from Noel Hidalgo of BetaNYC. This year, BetaNYC and the fellows have been charged with building tools to make open data more useful to the unique needs of Manhattan's 12 Community Boards. I would especially like to thank Mary McCormick from the Fund for the City of New York for making this program possible.

The Open Data Law has brought the civic hacker movement to the forefront of good government oversight. Open data has become more than just transparency; it makes government more accountable, serves as a teaching tool for undergraduates and a pipeline to good jobs, empowers citizens and small businesses, and improves city services. There is much progress to be made on this exciting front, and I will do all I can to ensure that the vision of 2012 continues to expand these initiatives, and that New York City remains a national leader in municipal data innovation.

To: NYC Council - Committee on Technology
From: Noel Hidalgo, Executive Director of BetaNYC

Re: Written testimony for NYC's open data laws oversight hearing

Wednesday, 21 September 2016

To the NYC Committee on Technology & Chairperson Vacca,

Introduction.

BetaNYC is a civic organization dedicated to improving all lives in New York through civic design, technology, and data. We envision an informed and empowered public that can leverage civic design, technology, and data to hold government accountable, and improve their economic opportunity.

We were founded in 2008 as a "meetup" to discuss open government in NYC. Our work empowers individuals and local communities to build a civically-engaged technology ecosystem and provide for an honest and inclusive government. We want New York's governments to work for the people, by the people, for the 21st century.

BetaNYC demystifies design, technology, and data to the point where anyone can use it, create it, and participate in the decision making process. We host a number of online platforms that provide the general public a mechanism to share ideas and data.

In the last twelve months, our community has grown 700 new members. We are now over 3,700 civic hackers who are ready to use our talents to help our neighbors.

How open data has grown our community.

In the last twelve months, we have hosted four significant events—NYC School of Data, NYC TreesCount Data Jam, a BetaTalk conversation on affordable housing data, and another BetaTalk on the release of the City's second largest data set, NYC 311 call inquiry data. We feel that these events have shaped where we are today, and point to a clear future.

We would like to thank Council Speaker Melissa Mark-Viverito; Manhattan Borough President Gale A. Brewer; Council Member Ben Kallos, NYC Department of Information Technology and Telecommunication; NYC Parks; NYC 311; Mayor's Office of Data Analytics; and the Mayor's Office of Technology and Innovation for joining us as partners.

We would also like to thank all of the government employees who joined us for all of our events. Without their expertise, we would never properly demystify how government technology and data works.

How we use the City's open data.

Re: Seven new pieces of open data legislation.

We fundamentally believe that these seven pieces of legislation were the right additions to make. We can see that they have strengthened the City's open data practice.

We were honored to be a partner with NYC Parks and prototype examples of good data dictionaries and geospatial / address standards.

We commend the Mayor's Office of Data analytics (MODA) for engaging with our community to gather public feedback and help ensure that the City's data users have a voice.

We absolutely agree with our colleagues at Reinvent Albany and NYPIRG on the following:

- *automating 100 datasets in 2016. That makes a total of 200 datasets out of 1,600 automatically updating on the Open Data Portal whenever the agencies internal data changes.*
- *publishing important new datasets including the City Budget, City Record Online, Seven Major Felony crime data, a huge TLC trip dataset, and NYC 311's call inquiry data.*
- *thirty of eighty agencies reported on FOIL responses that included public data that is or should be on the Open Data portal.*
- *DoITT and MODA staff are reading comments and requests on the Open Data Portal and responding.*
- *The administration's Open Data Team published their annual update on-time.*

Re: NYC's Civic Innovation Fellows with Manhattan Borough President Gale A. Brewer

Reference URL < https://beta.nyc/programs/nyc-civic-innovation-fellows/ >

Since July 2015, we have partnered with the Manhattan Borough President Gale A. Brewer to develop an innovative civic engagement program we call the Civic Innovation Fellows. With financial support from the Fund for the City of New York and Data and Society Research Institution, we have explored how Community Board district offices can better use data and technology.

We have reviewed Manhattan Community Board offices, their data & analytics capacity, and how they share information across digital streams. We discovered that Community Boards have a desire to use open data but don't always have the bandwith, education, nor tools to process the City's data. Frustratingly, we discovered that there are zero best practices bringing Community Boards to use 21st century tools nor teaching them how to use NYC's open data.

Using insights developed with Pratt Institute's SAVI program, we have outlined a framework to teach Community Boards, Council offices, and community based organizations how to use NYC 311's service request data and NYC Park's TreesCount data. Frustratingly, we have not found any financial support to teach the City's data to its residents.

NYC 311's service request and NYC Park's TreesCount data set are two of the 1,500 datasets that should have educational rubrics attached to their data dictionaries. **While we love the data dictionaries law, this should be considered the floor, not the ceiling. Every data dictionary should contain mini-tutorials explaining how to best explore the data and how to embrace the data portal's functions.**

While some of my colleagues might criticize the current data portal's user experience, a skilled user can quickly navigate around a dataset and easily produce reports. Granted, you will need a lightning fast connection, a large monitor, fast computer, and a bit of luck, but for now, we have the best portal tool for the widest audience. **Moving forward, NYC needs to develop tools to better suit bulk data users. We encourage the city to explore open source data sharing tools that will give agencies the flexibility to host and share their own data. Fundamentally, we want Agencies accountable for producing high quality datasets.**

This year, will be working with the Manhattan Borough President and the Fund for the City of New York to improve our Civic Innovation Fellows program and create a simplified online curriculum for all. **We should note that this is something that this Administration and the Council would benefit from and we hope to have your investment as we build the next level of open data education.**

Re: NYC Parks Data Jam

Referance URL < https://beta.nyc/2016/08/05/treescount-data-jam-2016-report-back/ >

The most important insight from the NYC Park Data Jam came from the evolution of a "*user centered data release workflow.*" Building off of user centered ideas, we believe that data sets should go through some user testing and like all technology products fit into a continuous improvement loop.

When is comes to continuous improvement of NYC's most valuable, most used data sets, we believe that every agency and every dataset should go through this release workflow.

- **Phase 1: Research & Discovery**
  - Establish target audiences
  - Draft data standards that appeal to broadest possible audience
  - Draft data dictionary
- **Phase 2: User Testing**
  - Release sample dataset for feedback
  - Perform user testing and get feedback on data dictionary & dataset
  - Develop a framework and/or guide as how to explore the dataset's important values
  - Allow time for revisions
- **Phase 3: Initial Deployment**
  - Upload data set and data dictionary to Portal with an event or video explaining the key features of the data.
  - Allow time for user testing on Portal and gather open feedback.
  - Share insights and communicate them out to the public. Continue to gather public feedback.

NYC Parks Data Jam Event Metrics
- Total Participants: 196

- Data Jammer Participants: 144
- Workshop Participants: 27
- Participant Diversity: 45% male / 39% female / 16% unspecified / 1% other
- Community Group / Public Stakeholder Groups: 5
- Children in on-site daylong childcare: 7
- Projects built: 23
- Winners: 5
- Organizers: NYC Parks and BetaNYC
- Partners: NYC Mayor's Office of Technology and Innovation, NYC Open Data, Civic Hall, Microsoft, Carto (was CartoDB).

Re: NYC School of Data

Refernace URL < https://schoolofdata.nyc/a-brief-recap-of-nyc-school-of-data-2016/ >

NYC School of Data proved there is a massive community in NYC that wants data and technology to be demystified. For one day, we featured 18 sessions, 40 presenters, 16 came from NYC Government, & three were elected officials.

A quarter of the event contained a data jam looking at how to address economic justice.

A critical component to our event was child care. This ensured diverse attendance. Several of our speakers couldn't attend if we didn't offer childcare. For older children, make it explicit they could attend with their parents.

Should the city host an open data summit, please offer child care. It makes a fundamental difference.

NYC School of Data event metrics
- We featured 18 sessions, 40 presenters, 22 are women, 16 came from NYC Government, & three are elected officials.
- 372 tickets; checked in 260+ people; 11 children under 10 y.o. attended the event.
- 49 people offered to volunteer. 30 labored.
- Eight NYC High Schoolers learned to navigate NYC's open data portal, manipulate NYC's 311 data, & map this data into CartoDB.
- One family attended and represented three generations: Grandma—organizer and data visualizer, Father—data advocate, & Daughter—aspiring civic hacker.
- Our event leadership team consisted of seven, three are women & five are people of color.
- For a third of the day, our hashtag was a trending topic on Twitter — #nycSoData.
- Organizer and Host: BetaNYC
- Partners: Manhattan Borough President Gale A. Brewer, Fund for the City of New York, Data and Society Research Institution, Microsoft, Carto (was CartoDB), Accela, Code for America, Internet Society of New York, NYC Mayor's Office of Technology and Innovation, NYC Mayor's Office of Data Analytics, Council Member Ben Kallos, Reinvent Albany, and Civic Hall.

Insights from the last year

Three types of data users.
While there are many types of data users, this year's research has led us to see three general types of data users.
- The general public who wants to see what the numbers mean.
- The data hacker who wants to see the data and play with the data.
- The scientist, business, or government entity who wants to enter the matrix and download data bulk.

For NYC's open data practice to be the best in the world, it must consider that these are the three types of users and each one has their own unique needs.

Concern about the seven new open data laws.
The passage of these seven new laws demonstrates that NYC's open data program has hit maturity. As a municipal practice, the city needs an investment ensuring this open data practice lives beyond the 2018 goal. Currently, the seven new pieces of legislation have created an unstable situation that will overwhelming the current open data team's ability to scale out NYC's open data practice.

Discrepancy in open data production.
We see several agencies who completely understand the benefit of open data and collaborating with the public. Yet we some who seemingly refuse to accept that law is an essential part of the 21st century.

In the past year, NYC 311, NYC Parks, Department of City Planning, NYC DoITT's GIS division, and the Taxi and Limousine Commission have produced data tools and data that are exemplary of the City's future. We hope their data teams continue to be supported and given the resources to lead by example.

More civic engagement events.
For the agencies that understand the value of open data, they have a unique opportunity to partner with the public and use hackathons and/or data jams to explore new insights and improve their data quality.

Quality data and data guides.
Community Boards and the general public are desperate for usable data. More importantly, they are desperate for content that will help them make sense of the data, aka an open data curriculum. San Francisco has a series of video guides that teach the general public on how to use their data portal < https://data.sfgov.org/videos >. We would love to see these videos appended to data dictionaries and see other mini-tutorials explaining how to best use the data portal's functions and how best explore the data.

Growing municipal data standards.
The future of municipal open data does not have one central clearing house processing all of the City's data and share it with the public. The internet does not work where there is one website servicing all of the world's information. While we agree that there should be one central catalogue of data resources, we fundamentally feel that Agencies need to own the responsibility in producing high quality datasets and understanding how the public uses its data.

For NYC to be the number one open data practice in the world, it must adopt a practice of establishing data standards, protocols, and coach Agencies to use those standards and protocols. NYC should be exploring

open source data sharing tools that give Agencies the flexibility to host their own data and interface directly with the public.

User center data release workflow.
When is comes to improving NYC's most valuable or most used data sets, we believe that datasets should go through a review process that bakes in public comment. This workflow is modeled after practices used in manufacturing and software development. We call this a user centered data release workflow. This workflow ensures that continuous improvement through public feedback will strengthen data quality and data products.

Enlisting public engagement through the data release process is a key part of every insight previously stated.

What we see'd like to see.

Fundamentally, you cannot cure NYC's open data future with a particular product but a dedicated practice.

We continue to encourage this Administration and the Council to place a significant investment in a dedicated open data team. This is the only way to ensure the City's data practice can scale across all City's agencies.

In an ideal world, we would like to see a Chief Data Officer resourced with a team and dedicated resources to shepherd the city's data practice, data standards, and outline education best practices. We do not see this office as counter to the current leadership. We see this team as a complementary—a policy, consulting, and product shop. From our research, this office would focus on the following.

Leadership & Standards
- With the City's technology and data leadership, draw together agency leaders to ensure data and technology standards produce data as a renewable resource.
- That technology procurement practices selects systems that allow for data version controls, API driven backends, with modular or open sourced capabilities.
- That these systems are always available, reliable, consistent, accessible, secure, and flexible to support an agency's mission.
- Help agencies perfect a feedback loop around data quality and public comments.
- Provide leadership to steer the production of citywide data policies and data standards.
- Ensure public feedback baked into the City's Technical Standards Manual (TSM).

Technology & Tools
- Prioritizing large scale data projects in conjunction with data owners.
- Standardize and automate future dataset dissemination.
- Sets policies for responsible data systems.
- Ensures that data reform and modernization—how agencies collect, uses, manages, and shares data—moves toward the stated goal of building "fact-based, data-driven decision-making" programs or policies.

- Oversees the development and stewardship of data sharing tools that enable agencies to share their own data and collect feedback directly from data users.

Evangelism

- The evangelism pillar would reach out to the public, industry, academics, and other branches of government to promote data, data services and tools.
- Inward, this pillar would connect the city's data managers to world's best and the brightest to ensure that Agencies are thinking about their best practices.
- Outward, this pillar would develop collaborations that further development of open data products and services.

Education & Trainings

- Inward, trainings would cut across agencies teaching and promoting data proficiency skills, geographic information systems, and data tools like MS excel, and R.
- Outward, this pillar would ensure that specific uses of data are taught and shared across the city. Community boards, institutions, and organizations would serve as a feedback loop to thoroughly define the public's need for technology, data, and data quality.

While some of these things can be done from the outside, we are at a make or break legacy point. For the City's open data practice to lead the world, they need to be done from within government.

Thank you for giving us this opportunity to give our story.

The BetaNYC Leadership Team

Dear members of the New York City Council Technology Panel,

We are Theo and Dalcini, and we are members of a newly created ad-hoc advocacy group related to privacy, specifically relating to online privacy.

The majority of our members are Open Data, Open Software, Freedom Software, Free Software and Transparency advocates who advocate for the sharing of all NYC citizens' data collected by private enterprise though government contracted programs like LinkNYC.

We are here to raise awareness of the problem that massive amounts of public data collection can do to privacy of NYC citizens. We have started our ad-hock advocacy group with the purpose of creating an official "New York City Privacy Guidelines Board."

The board would be composed of New Yorkers from various fields, including communications technology, security and law experts, as well as citizens, who would debate privacy implications and make recommendations to elected representatives when they discuss technology contracts that include the collection of private data, it's safe guarding and potential implications.

https://www.facebook.com/NYCPrivBoardAdvocates

Open Data Oversight Hearing - Sept 21, 2016

Your Honorable Council members,

I'm here to testify and petition the City Council on the need to modernize the Open Data infrastructure of the City.

New York City is a city like no other. I won't use your valuable time to go through a long list of reasons, but I believe we can all agree that one primary reason why its so is that its the global center of so many industries - Finance, Media, Fashion, Information Technology, and Advertising, to name a few.

Underpinning all these industries is the fact that they are all knowledge-based, services-driven, data-intensive professions - the Information and Knowledge Economy.

And if New York City is to become the epicenter of the emerging Urban Informatics industry - covering GovTech, CivicTech and Smart Cities, I further submit that the current Open Data Portal needs to be revamped to support this nascent industry.

For Open Data is Essential Infrastructure.

And as such, it needs to be (1) Open Source, (2) be a true Platform that can support innovation and experimentation, and (3) permit true Federation so that not only City Agencies can participate in growing this knowledge corpus, but third parties like universities, non-governmental organizations and the private sector can effectively collaborate with the City in creating a robust Data Commons open to public-private partnership.

Already, New York City is home to Civic Hall, BetaNYC and NYCBigApps. Its also home to two of the leading Open Source Data Portal providers in the world - NuCivic, now part of GovDelivery; and yours truly from Ontodia - a direct product of NYCBigApps, now part of OpenGov. Both companies are now working around the world helping governments at every level use modern Open Source, Open Data platforms.

An Open Source solution allows the City to (1) build internal capacity, (2) tap the genius of the community, and (3) accelerate innovation and compliance to the recent amendments to the City's Open Data Law.

As I'm sure you're keenly aware - being an urban innovation leader and an early adopter of technologies like Open Data, often comes at the cost of creating legacy investments, that quickly get leapfrogged, especially so in the technology industry.

I submit that the current solution is in dire need of a revamp, and the City should consider stopping the sole-source open data contract, especially now that there are multiple NYC-based providers in the space.

Thank you so much for your kind attention,
Joel Natividad

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____

☑ in favor ☐ in opposition

Date: 9/21/2016

(PLEASE PRINT)

Name: Aleksey Bilogur

Address: 2900 Ocean Ave.

I represent: Myself

Address: 2900 Ocean Ave.

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor ☐ in opposition

Date: _____

(PLEASE PRINT)

Name: JOEL NATIVIDAD

Address: 33 IRVING PLACE

I represent: OPEN GOV

Address: _____

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor ☐ in opposition

Date: 21 SEPT 2016

(PLEASE PRINT)

Name: NOEL HIDALGO

Address: 45 DRIGGS AVE, BKLYN 11222

I represent: BETA NYC

Address: _____

Please complete this card and return to the Sergeant-at-Arms

THE COUNCIL
THE CITY OF NEW YORK

$$\boxed{\textit{Appearance Card}}$$

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor ☐ in opposition

Date: _____

(PLEASE PRINT)

Name: John Kaehny

Address: _____

I represent: Reinvent Albany, NYPIRG, Common Cause NY

Address: _____

▶ *Please complete this card and return to the Sergeant-at-Arms* ◀

THE COUNCIL
THE CITY OF NEW YORK

$$\boxed{\textit{Appearance Card}}$$

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor ☐ in opposition

Date: 9/21/16

(PLEASE PRINT)

Name: ALBERT WEBBER

Address: Civic Engagement Manager - DoITT

I represent: _____

Address: _____

▶ *Please complete this card and return to the Sergeant-at-Arms* ◀

THE COUNCIL
THE CITY OF NEW YORK

| Appearance Card |

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor ☐ in opposition

Date: 9-21 16

(PLEASE PRINT)

Name: THEO + DALLINI

Address: _____

I represent: NYC PRIVACY BOARD ADVOCATES

Address: _____

➤ Please complete this card and return to the Sergeant-at-Arms ◀

THE COUNCIL
THE CITY OF NEW YORK

| Appearance Card |

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor ☐ in opposition

Date: 9/21/16

(PLEASE PRINT)

Name: Dr. Amen Ra Mashariki

Address: Chief Analytics Officer / Chief Open Platform officer

I represent: _____

Address: Director of Mayor's Office of Data Analytics

➤ Please complete this card and return to the Sergeant-at-Arms ◀