

CITY COUNCIL
CITY OF NEW YORK

----- X

TRANSCRIPT OF THE MINUTES

Of the

COMMITTEE ON TECHNOLOGY

----- X

September 21, 2016

Start: 1:00 p.m.

Recess: 2:35 p.m.

HELD AT: 250 Broadway - Committee Rm
14th Fl

B E F O R E: JAMES VACCA
Chairperson

COUNCIL MEMBERS: Annabel Palma
David G. Greenfield
Barry S. Grodenchik
Joseph C. Borelli

A P P E A R A N C E S (CONTINUED)

Dr. Amem Ra Mashariki
Chief Analytics Officer and Director
Mayor's Office of Data and Analytics

Albert Webber
Department of Information, Technology and
Telecommunication

John Kaehny, Executive Director
Reinvent Albany
Co-Chair, NYC Transparency Working Group

Noel Hidalgo, Executive Director
BetaNYC

Joel Trinidad
Open Data Provider

Tio Chino and Daphne
Data Privacy Advocacy Group

Alexa Billier, Student
CUNY Baruch College

2 [sound check, pause]

3 CHAIRPERSON VACCA: But it on vibrate so
4 we don't—we won't be hearing extraneous noise.

5 There's a lot of extraneous noise in these hearings

6 to begin with I want you to know. [background

7 comments] [laughs] I couldn't use this stuff. Okay.

8 Thank you. It's 1:00 p.m. and today is September

9 21st and I want to welcome you all to the Technology

10 Committee hearing here at the New York City Council.

11 I'm James Vacca. I'm chair of the Committee on

12 Technology, and I'm sure colleagues of mine will be

13 joining me throughout the hearing. We have one

14 colleague here, Council Member Barry Grodenchik from

15 Queens, and I welcome him and other committee members

16 that will be arriving. We're here today to conduct

17 an oversight hearing on the Open Data Law, the 2016

18 Open Data Plan and the seven amendments to the Open

19 Data Law that were passed by this committee within

20 the past year. I sponsored two laws in that package,

21 Local Law 110 of 2015, which requires the timely

22 updating of certain public data sets on the Open Data

23 Portal, Local Law 8 of 2016, which requires citizen

24 examinations and verifications of the compliance of

25 certain agencies with the requirement to public data

2 sets. Additionally, this committee passed five other
3 bills, Local Law 106 of 2015 sponsored by Council
4 Member Cabrera, which requires the preservation and
5 archiving of data that would have otherwise been
6 purged from the Open Data Portal. Local Law 107
7 sponsored by Council Member Gentile, which requires
8 the creation of a data dictionary for every data set.
9 Local Law 108 of 2015, sponsored by Council Member
10 Ben Kallos, which standardized address data, a geo
11 code address data. Local Code 7 of 2015 sponsored by
12 Council Member Anabel Palma, which requires agencies
13 to review FOIL requests that include the release of
14 data for inclusion of that data on the Open Data
15 Portal, and Local Law 109 of 2015, sponsored by
16 Council Member Torres, which creates the response
17 timelines for public requests on the Open Data
18 Portal. Today, we will discussing the implementation
19 of these seven local laws, the challenges that have
20 arisen, the ongoing issues, and the ways we can work
21 together to solve them both administratively and
22 legislatively. The Open Data Law passed in 2012
23 requires DOITT to work with all city agencies to post
24 public data online in a centrally accessible
25 location, the Open Data Portal by 2018. As mandated

2 in the Open Data Law, DOITT must annually release an
3 Open Data plan. The 2016 plan was released on July
4 15, and reports that the portal now contains over
5 1,500 data sets, over 200 of which were automated
6 with an additional 212 data sets planned for release
7 before 2018. Before we talk about the ways to
8 improve the Open Data Law, I want to commend DOITT
9 and the Mayor's Office of Data Analytics for their
10 dedication to the implementation of this law. Like
11 on so many other fronts, New York city is a
12 nationwide leader on open data, and the hard work of
13 DOITT and MODA along with help from the Council and
14 dedication of advocates certainly deserves
15 recognition. Yet, there is still work to be done,
16 and the committee intends to advocate for ongoing
17 improvements. This committee believes that data
18 quality, and the ability to report errors found
19 within the data have become a serious concern, and
20 needs to be addressed. This committee also remains
21 concerned with agency compliance both with the
22 original law's mandates, and the mandates from the
23 recent amendment made by the Council. The number of
24 agencies reporting on their FOIL data releases as
25 required by Local Law 7 were disappointingly low.

2 I'm hoping that today's oversight hearing will shed
3 light on the successes of open data as well as on the
4 many challenges we still face. With that said, I
5 want to mention I've been joined by my colleagues.
6 To my left Council Member Joseph Borelli. He's
7 usually always on my right, but I'll take him on my
8 left. This is actually on my right, and I will turn
9 it over to Dr. Amem Ra Mashariki of the Mayor's
10 Office of Data Analytics and Albert Webber of the
11 Department of Information, Technology and
12 Telecommunication, and we look forward to hearing
13 your testimony. I do have to swear you in. Do you
14 affirm to tell the truth, the whole truth and nothing
15 but the truth in your testimony before this
16 committee, and to respond honestly to council
17 members' questions?

18 DR. AMEN RA MASHARIKI: Yes.

19 ALBERT WEBBER: Yes.

20 CHAIRPERSON VACCA: Okay, so thank you.

21 Who would like to begin?

22 DR. AMEN RA MASHARIKI: I will, Chairman
23 Vacca. Thank you.

24 CHAIRPERSON VACCA: Introduce yourself,
25 please.

2 DR. AMEN RA MASHARIKI: My name is Dr.
3 Amen Ra Mashariki. I'm New York City's Chief
4 Analytics Officer as well as the Director of the
5 Mayor's Office of Data and Analytics. I wanted to
6 start by saying thank you very much for inviting us
7 here to talk about the incredible strides that we're
8 making—we've made over the last year with the work on
9 open data. Again, my name is Dr. Amen Ra Mashariki.
10 I'm the Chief Analytics Officer and Chief—and also
11 Chief Open Platform Officer for the City of New York.
12 I'm joined by Albert Webber, of the Department of
13 Information, Technology and Telecommunication. We're
14 here today to speak about the way MODA and DOITT have
15 worked together over the last year to fulfill the
16 requirements of the city's Open Data Law and its
17 recent updates. With respects to open data and open
18 data for all, when city leaders passed Local Law 11
19 of 2012, New York City set a high bar for effective
20 and transparent government. Not only would we open
21 our data, we would open it all. Last July, we
22 committed to open data for all, a more inclusive
23 vision in which all New Yorkers, not just the tech
24 savvy can find value in all the data. We believe
25 that every New Yorker can benefit from open data, and

2 open data can benefit from every New Yorker. Soon
3 after we published *Open Data for All*, Council passed
4 and Mayor de Blasio signed a package of bills
5 bolstering the original open data law with
6 requirements to make it easier for New Yorkers to
7 access, understand and derive value from city data
8 online. These new laws required us to be more
9 responsive to public feedback, create better
10 technical standards and geospatial data, and data
11 retention, and strengthen our means of ensuring
12 agency compliance. Together, they help anchor the
13 Administration's commitment to transparency and
14 equitable uses of technology around open data. One
15 of the things that we've been able to do is build and
16 open data ecosystem. MODA is the business owner of
17 this initiative. As the hub of analytics in the
18 city, MODA advocates for the use of open data in
19 citywide data analytics. DOITT is the technical
20 owner ensuring that technological capabilities are
21 constantly improving to better meet user's needs.
22 Agencies are the data owner. Our primary points of
23 contact with agencies are their open data
24 coordinating with whom we work closely to ensure that
25 agencies comply with the law. These three entities

2 along with our friends in the City Council and the
3 civic tech community help to form the city's Open
4 Data Team. Together, we continue to improve the
5 biggest and best municipal open data initiative in
6 the world by focusing on one number, 8.4 million, the
7 total number of New Yorkers. Over the last year, the
8 new statutory mandates have catalyzed and alignment
9 of our vision of open data for all with our technical
10 tools and work processes. In order to implement
11 these laws while growing, the usability, reach and
12 impact of open data, we have drilled down on the
13 following three areas: User research, feedback
14 mechanisms and technical standards. With respects to
15 user research, the first step in opening data should
16 be focusing on demand. We support creative analytic
17 thinking to enable our users to answer the question
18 most important to them, but we also hope to spur
19 demand in communities less familiar to open data.
20 Open data for all means open data for app developers,
21 agency analysts and mapping gurus, but also open data
22 for Emma Yang, the 12-year-old New Yorker who
23 published her analysis of NYPD data from the Open
24 Data Portal, as the youngest ever recipient of
25 Wolfburn's Research Mentorship program. To this end,

2 MODA has partnered with the academic community to
3 examine where open data is reaching New Yorkers and
4 where it's not. A Capstone at the New York
5 University Center for Urban Science and Progress
6 measured data poverty across the city. So we
7 partnered with NYU CUSP on a Capstone project where
8 we measured data poverty across the city. Data
9 poverty refers to a condition in which communities or
10 people lack access to, use of or representation
11 within data that is nevertheless used to inform
12 decisions that may affect them. This analysis is
13 among the first of its kind, and will be released
14 publicly later this fall. We also did Capstone with
15 Columbia SIPA. We work with the Columbia University
16 School of International and Public Affairs to assess
17 existing and future opportunities for open data to
18 advance the goals of the city's community base
19 organization. Although municipal data is in high
20 demand, CBOs pointed to several pain points in
21 accessing and operationalizing open data. These
22 results are informing conversations on ways to
23 improve user experience on the portal. Local Law 107
24 of 2015 is helping ensure that data is more
25 understandable for non-technical users by requiring

2 that every data set include a data dictionary
3 defining the attributes and clarifying the usability
4 of its contents. With respect to feedback
5 mechanisms, the Open Data Portal is not a static
6 product. It is a platform that should be used as a
7 tool for public feedback. We are dedicated to
8 soliciting user input at every step of the open data
9 life cycle to improve overall data quality and
10 increase data use. We want those who give feedback
11 to understand what their contributions are being used
12 for, what opportunities for change are available, and
13 when they can expect to see the changes take effect.
14 Substantive engagement is more than obligation. It
15 must be default for open data. We did a comment
16 analysis as well. This summer we reviewed every
17 comment that has ever been submitted to the Open Data
18 Portal, and developed a typology for the types of
19 inquiries that come in. This has informed how we are
20 categorizing future feedback in order to optimize the
21 way we deliver customer service. Over the next year,
22 we will be implementing a new technological solution
23 to intake feedback with a centralized mechanism that
24 allows for quicker responses and better tracking on
25 our interactions with users. This will make it

2 easier for us to meet the requirements of Local Law
3 109 of 2015, which mandates timely responses to
4 public requests. A better feedback mechanism gives
5 us a better way to comply with the law. This spring
6 as per Local Law 8 of 2016 on examination and
7 verification, we worked with the Department of
8 Investigations to develop a plan for examining the
9 Housing and Preservation Department, the Department
10 of Sanitation, and the Department of Correction for
11 information assets that may contain public data.
12 Before this year, we would work with agencies to
13 craft publishing plans for their data. Now, the
14 Examination and Verification Law, the Public Request
15 Law and the law requiring coordination between open
16 data coordinators and portal officers have enabled
17 new means of locating data covered by the original
18 Open Data Law. They also help enact a broader
19 culture shift. As open data becomes the norm for
20 city data, it makes agencies more aware of the data
21 they have, and the data they produce and spurs better
22 upkeep and disclosure of information. Many of the
23 new pieces of legislation require updates to the
24 Technical Standards Manual, the document of record
25 that outlines all policies and protocols for open

2 data. As we evaluate each standard, we have three
3 priorities, meeting the letter of the law, meeting
4 the spirit of the law and maintaining usability to
5 the maximum number of users. To do this effectively,
6 public participation is vital. Take for example
7 Local Law 108 of 2015, which mandates the formation
8 of a working group to create standards for address
9 and geospatial information on the Open Data Portal.
10 Over the last several months, I convened a group of
11 Open Data coordinators to draft recommendation to
12 geospatial attributes, column headers and geocoding
13 processes for data sets. The recommended geospatial
14 standards reflect that data builds (sic) most
15 frequently captured by city agencies, information
16 that is in highest demand from public users, and
17 attributes that will have the biggest impact on
18 citywide operations once they are standardized. In
19 the spirit of open data for all, we have invited the
20 public to join the conversation. We received
21 constructive feedback and helpful suggestions from
22 members of the tech startups, national non-profits,
23 research uni--universities and others. For example,
24 once submission from the American Society for the
25 Prevention of Cruelty to Animals, ASPCA, said, "These

2 new standards are a great step in the right direction
3 and will allow non-profits like ours to avoid costly
4 geocoding of addresses.” These comments are helping
5 us refine our recommendations to the standard, which
6 will ultimately be included in an update—in an
7 updated version of the Technical Standards Manual
8 later this year. I’d like to close by sharing a
9 standout account of enormous potential of open data.
10 This spring, Ben Wellington, a local data scientist
11 who studied parking violation data from the
12 Department of Finance, and found that some of the
13 locations where cars were ticketed were, in fact,
14 legal parking zones. When he brought his analysis to
15 the attention of the—of the city, NYPD analysts
16 confirmed his results also using open data, and took
17 steps to ensure that ticketing officers would avoid
18 the oversight in the future. Wellington, who runs
19 the popular analytic blog, Icon NY, wrote the
20 following when he saw PD’s response: “I was
21 speechless. This is what the future of government
22 could look like one day. This is what open data is
23 all about. Imagine a city where all agencies embrace
24 this sort of analysis instead deflect and hide from
25 it. Democracies provide pathways to government to

2 learn from their citizens. Open data makes those
3 pathways so much more powerful. In this case, NYPD
4 acknowledged the mistake, is retraining its officer
5 and is putting and monitoring to limit this type of
6 erroneous ticketing from happening in the future. In
7 doing so, they have shown that they are ready and
8 willing to work with the people of the city, and what
9 better gift can we get from open data than that. Open
10 data lies at the nexus of digital services, public
11 transparency and cutting edge civic analytics. But
12 we are constantly aiming higher. We will continue to
13 test our assumption, and try new ideas, collect
14 information on our efforts and analyze our
15 performance for points of improvement. We have
16 committed to engaging the public throughout the
17 process to keep us honest and on track as we continue
18 to learn and improve. I would like to thank the
19 Council Technology Committee for the opportunity to
20 testify today, for its continued support of open
21 data. At this point, I would like to turn it over to
22 Albert Webber, who will walk you through our progress
23 on each piece of legislation individually in order.

24 ALBERT WEBBER: Thank you, Dr. Mashariki,
25 and good afternoon, Chairman Vacca and members of the

2 City Council Committee on Technology. My name is
3 Albert Webber, and I will be testifying today on
4 behalf of the Department of Information Technology
5 and Telecommunications. As you know, this year's
6 annual update to the NYC Open Data Plan was released
7 on July 15, 2015. In this plan, DOITT with the
8 Mayor's Office of Data Analytics and MODA and city
9 agencies collaborated to increase the quantity and
10 quality of data in the NYC Open Data Portal. Since
11 last year's update traffic on the Open Data Portal
12 has increased receiving over five million hits with
13 more than 2,000 user created use stream barriers.
14 (sic) To date, nearly 1,600 data sets (sic) have
15 been made available, 116 of which were released in
16 the past year as per the Open Data Plan. These
17 include restaurant and building inspections, the city
18 record online, and the City Council own Constituent
19 services. The infrastructure has been built for the
20 automation of over 100 new datasets bringing the
21 total number of automated requested data sets to over
22 200. Additionally, we have identified and published
23 more than 40 unscheduled data sets. This
24 demonstrates the commitment of agencies to make data
25 available as quickly as possible. To continue

2 building on this past year's accomplishments, DOITT
3 is actively seeking to expand its Open Data Team. We
4 are currently interviewing candidates for three or
5 four new positions that will be fully dedicated to
6 the Open Data program. With this additional staff,
7 we plan to improve among other things our
8 communication the open data users, our technical
9 infrastructure, and the quality of receiving data.
10 The package of laws the City Council passed late last
11 year and early this year is an integral part of open
12 data retention. (sic) The following portion of
13 testimony will outline our progress of each law.

14 Local Law 106 of 2015 with regard to the
15 retention. This law requires DOITT to preserve
16 record level data and to establish guidelines for the
17 division of large data sets into smaller ones. The
18 new guidelines will facilitate archiving of
19 historical data as data sets get updated, allow users
20 to follow trends over time, and improve the overall
21 user experience. We are ahead of schedule for
22 complying with this legislation. We have already
23 begun working with agencies so data is preserved to
24 maintain accurate and active lists for some of our
25 more popular data sets. Appending instead of a place

2 in the existing records, and dividing some of the
3 larger data sets into smaller, more manageable data
4 sets. The Open Data Team has already drafted the
5 standards for data retention and will open them up
6 for public feedback later this fall.

7 Local Law 107 of 2015 provides the data
8 dictionaries. This local law requires DOITT to
9 include a plain language data dictionary with every
10 data set on the Open Data Portal. The data
11 dictionary has made data more understandable by our
12 users, and fits into the vision of open data for all.
13 To help ensure that that data dictionaries have a
14 view to the public, we plan to collaborate with users
15 to create data dictionary templates, and to open our
16 open data—our open—our data dictionary standards for
17 public comment later this fall. Since the law was
18 enacted, we have—are already out of the box at least
19 70 data dictionaries, and we'll work towards
20 including data dictionaries for all data set in the
21 Open Data Portal by the end of 2017.

22 Local Law 108 will provide for geospatial
23 and address those needs. This law requires the
24 creation of standardized geospatial address layouts
25 for all data sets containing the address information.

2 This standard will result in more robust data being
3 added to the portal that allow users to easily
4 perform spatial analysis across the multiple
5 datasets, and build applications at levels with more
6 information. As previously mentioned, Mayor de
7 Blasio designated Dr. Amen Ra Mashariki Chief-Chief
8 Analytics Officer to continue a working group that
9 will provide recommendations of creating uniformed
10 geospatial standards to both DOITT and the City
11 Council. To develop these standards, we've held
12 several meetings with open data coordinators, agency
13 stakeholders, the Department of City Planning's Geo
14 Support team and open data advocates. We then
15 gathered this feedback, drafted standards, and opened
16 it up to the public earlier this month. We look
17 forward to finalizing these mandates and providing
18 the technical value to agencies to adhere to this
19 legislation.

20 Local Law 109 provides for the timely
21 responses to public requests. This law requires
22 DOITT to provide additional responses to requests for
23 public data sets within two weeks of receipt, and the
24 agency to which DOITT referred a request to make a
25 final determination within two months of receipt.

2 Over the past year, the Open Data Team has worked
3 closely with agencies to respond to data requests
4 dating as far back as 2011. We found that many of
5 these requests for data sets were—already existed,
6 reinforcing the notion that users have a true
7 knowledge of the city’s unique data landscape, and
8 the agencies have been proactive about their use of
9 data. Improving feedback has continued to be a prior
10 of the Open Data Team moving forward. In the coming
11 months, we will be improving and consolidating our
12 communication channels to make sure there’s a clear
13 line of communication between us and our user. We
14 invite the public to continue requesting data sets,
15 to ask us questions and to comment on our proposed
16 data standards.

17 Local Law 110 of 2015 provides the time
18 of the data set updates. This law requires data sets
19 updated on agency websites to also be updated on the
20 Open Data Portal and fits with our strategy to
21 provide regularly and time requests on accurate data.
22 For timely updates, our team will work with agencies
23 to ensure there’s a schedule in place to maintain up-
24 to-date data on the portal. Regarding manually
25 updated updates, our new staff will play a major role

2 in ensuring that data is refreshed, but will also
3 work to identify data and use of websites that will
4 require the annual updates. There are roughly 1,400
5 data sets that are not automatically updated, and we
6 will work to ensure that they are all addressed.

7 Additionally, over the last year, our technical team
8 has made huge strides in making timely data set
9 updates. For the first time we began the process of
10 automatically update geospatial data sets, a process
11 that will save us hours of work every year.

12 Additionally, partnering with the Department of
13 Sanitation our team will be releasing the raw data
14 source from the PlowNYC application, which has been
15 in high demand for many years. The data will be made
16 available in year real time than snow events with
17 plow locations being updated several times per hour,
18 and older records being archived. This is the first
19 time we will be able to deliver automated data, and
20 more frequently than daily, and we hope to apply
21 these technology enhancements to other high demand
22 data sets.

23 Local Law 7 of 2016 provides for FOIL.

24 This law requires agencies to review responses to
25 Freedom of Information Law requests that include

2 release of data to determine if such responses
3 received public data sets that have not yet been
4 included on the Open Data Portal. The results of
5 this review were included in the last update to the
6 Open Data Plan, and will be included in updates going
7 forward. To facilitate agency compliance through its
8 Open Data and legal staff provided guidance to agency
9 oversees and counsel's office. It fully laid out
10 what agencies needed to consider and was reading that
11 call with us.(sic) This legislation is extremely
12 helpful in expanding open data discussions within
13 agencies, and highlighting the association between
14 data that is requested through FOIL and open data
15 both of which are indicative of promoting
16 transparency. Reporting on these metrics included
17 input from open data coordinators, FOIL officers,
18 general counsel, and communications officers, and we
19 anticipate they will continue to help agencies
20 prioritize data going forward.

21 Local Law 8 of 2016 with regard to the
22 examination and verification process. So, last, but
23 certainly not least, Local 8 of 2016 require the
24 Mayor to appoint an office or agency to conduct a
25 series of examinations and verifications of

2 individual agencies' compliance with the Open Data
3 Law. MODA is that appointed office, and began the
4 process of collaborating with the Department of
5 Investigation and its commissioner on examination and
6 verification compliance. All agencies that must
7 comply this year, the Department of Housing
8 Preservation and Development, the Department of
9 Corrections and the Department of Sanitation have
10 been briefed on the examination and verification
11 schedule, and MODA was to give surveys to each agency
12 by the end of September. Additionally, suggestions
13 from the public will be collected throughout the
14 fall. Thank you for the opportunity to testify this
15 afternoon. These have succeeded in making the Open
16 Data Law stronger, and we have been pleased to update
17 on our progress. Open data remains a priority of
18 this Administration and we thank our partners in the
19 City Council and the civic community for their
20 continued advocacy. And this completes our prepared
21 testimony, and we look forward to answering any
22 questions.

23 CHAIRPERSON VACCA: I want to thank you.

24 I want to note we've been joined by Council Member
25 Annabel Palma to my right. Your testimony was good.

2 I-I want you to know I'm pretty happy with one thing,
3 and that is the Plow New York City action taken.
4 Last year we had complaints. I think we've made a
5 little bit of news here. It's not easy for this
6 committee to make news I want you to know.
7 [laughter] But we did make some news today because
8 winter is here, believe it or not, coming up on us,
9 and here we had a situation where last year the Plow
10 New York City information was telling us where these
11 spreaders were going, but we did not retain the
12 information of where they have plowed, and people are
13 saying that there's people who have never plowed, and
14 then when they went into open data, they could only
15 find that information going forth not going back. So
16 here we are told that, you're going to be address
17 this year, and plow locations will be updated several
18 times per hour, and all the records are being
19 archived. This is the first time this is happening,
20 and I think this will give citizens a real handle on
21 whether or not their street has been plowed. I
22 cannot begin to tell you how many phone calls I get
23 from people telling me that their street was never
24 plowed during the heavy snow storm. How many times
25 those streets were plowed, but the reality was is

2 that it was plowed, but everybody dug themselves out.
3 They dug their cars out on the street, shoveled it
4 back, and if the temperatures don't reach 32 or
5 above, the snow has nowhere to go. Nothing melts. So
6 here we will have a record—Excuse me. So, here we
7 will have a record of that taking place, and I think
8 that's good. So I thank you. So that will be in
9 place for this year. Yes. Okay. What's the most
10 common reason people use open data? What are they
11 looking—what do you find that they're looking for
12 more frequently? I know when you call 311, the—the
13 main answer—the main question probably is noise. So
14 how frequent are the calls to open data about a
15 particular—are there particular things people are
16 looking?

17 ALBERT WEBBER: So reviewing the data
18 sets requested, combinations of communities, I'd say
19 that the primary request is the violations that we've
20 heard, violations from the Environmental Control
21 Board and violations for restaurants.

22 CHAIRPERSON VACCA: Okay.

23 ALBERT WEBBER: [off mic] So—so overall
24 it would be violations and additionally data from the
25 MCA, which actually doesn't fall under the

2 jurisdiction of the city of New York. The violations
3 and MTA data they're probably as much as has been
4 requested. (sic)

5 CHAIRPERSON VACCA: So people look for
6 MTA data but it's not there?

7 ALBERT WEBBER: [off mic] It is there.
8 We—we do link to data that the MTA makes available,
9 but we don't own or retain that data.

10 CHAIRPERSON VACCA: So what do people
11 look most frequently for? I'm sorry.

12 ALBERT WEBBER: Other than that,
13 violation. It's the most broad—

14 CHAIRPERSON VACCA: [interposing] You
15 mean from Environmental Control Board violation data?

16 ALBERT WEBBER: [off mic] I mean from a
17 variety of agencies. So whether it's violation
18 through restaurant inspections the Department of
19 Buildings and the Environmental Control Board. The
20 violations tend to be the most frequently accurate
21 data.

22 CHAIRPERSON VACCA: Now many of the
23 portals will be zeroed in on—zeroed in on for
24 updates? You said you're going to be providing
25 updates to all the information or this going to be

2 portal by portal? How will the updates be provided?
3 How frequently and where?

4 ALBERT WEBBER: With regards to all data
5 sets?

6 CHAIRPERSON VACCA: Yes.

7 ALBERT WEBBER: [off mic] So, it-it
8 depends on the updates that have been set by the
9 agency. We work with agencies to I guess recommend
10 appropriate frequencies that we think the public
11 would want. At the end of the day, the agency decides
12 within their business concerns what is the most
13 reasonable update frequencies. From our end, we're
14 trying to automate whatever it is that is mandated
15 that is possible. We have 200 automations right now,
16 but we're continuing to try to build that number, and
17 then hopefully the new staff that we'll bring in will
18 help us to also continue with that number.

19 CHAIRPERSON VACCA: Let me ask you so
20 more questions. What agencies have not been as
21 forthcoming as they should have been? I'm thinking
22 that the police don't—I'm trying to be nice, you
23 know, with politics-political-political terminology,
24 but I'm thinking of the Police Department, and I know
25 that we have gaps here, and I want to know what we're

2 doing. What agencies besides the police have been a
3 little behind their timetable, and what are these,
4 and what are the agencies number one? What have you
5 done to—what is the extent of their being behind, and
6 what have you done to get them to—to the table as
7 for—as for the law?

8 DR. AMEN RA MASHARIKI: So I think
9 agencies have made strides across the board. So with
10 regard to the NYPD, crime data has been something
11 that's been heavily requested over the last two
12 years, and—and I believe it was last year, they
13 actually released for the first time intimate level
14 crime data. So we can go to neighborhood for the
15 seven major felonies in New York City and actually
16 see and map out where these crimes are occurring.
17 NYPD has also added a number of data to their
18 actually website, which we also went to with a
19 variety of different data sets. So I would say that
20 the NYPD has made real strides going forward. I mean
21 overall, I—I wouldn't say any particular agency is
22 not fully committed to doing what they're supposed to
23 do. Agencies had progress across the board, and
24 we've had automation through—through all sorts of
25 agencies, and all sorts of new data sets as per

2 requested on the Open Data Portal. So I can't say
3 there any one set that they haven't been doing what
4 they need to be doing for the Open Data Law.

5 CHAIRPERSON VACCA: The Police
6 Department, they--has complied?

7 DR. AMEN RA MASHARIKI: They've--I mean
8 yes.

9 CHAIRPERSON VACCA: So they're--so I was
10 wrong. They--they have submitted everything on time,
11 and we have all the information on--on open data we're
12 supposed to have in the PV. (sic)

13 ALBERT WEBBER: So they've--so if we break
14 down each legislation one by one, I'd have to go back
15 and get actually exactly what we may not be using
16 that they've been using. More data set over the past
17 year. They've responded to the public data set
18 requests. I'd have to--I'd have to go back to any
19 specifics, but they've made real strides over the
20 last year.

21 CHAIRPERSON VACCA: Why--well, I'd--I'd
22 like you to get back to me, and I'm going to--when you
23 get back to me, I'll get back to you because there
24 are agencies we've analyzed that we don't think we
25 have total compliance with. I--I wanted to ask

2 further, though, about that. Can you meet the
3 deadlines set by the legislation in all cases that
4 the Council has legislated? Are you able to meet
5 those deadlines?

6 DR. AMEN RA MASHARIKI: It would be
7 difficult with the newly—the four lines that we're
8 going to bring on board. So with the four lines that
9 we're going to bring on board, we do feel confident
10 that we can hit the numbers with all the goals in
11 this legislation.

12 CHAIRPERSON VACCA: Let me—let me clarify
13 that. Are you interviewing right now, or are you
14 advertising?

15 ALBERT WEBBER: We're interviewing.

16 CHAIRPERSON VACCA: You're interviewing
17 so the advertisement phase is finished.

18 ALBERT WEBBER: Yes.

19 CHAIRPERSON VACCA: Why is that some
20 agencies now are hiring open data liaisons, and other
21 agencies are trying to use existing personnel with
22 previous responsibilities to be their open data
23 liaison? There is not a consistency? I feel every
24 agency at this point has to have a dedicated open
25 data person, and many of the agencies do not.

2 They're looking to use people that wear several hats,
3 and I think that that minimizes the importance of
4 what we're trying to do when it relates to
5 transparency and government accountability. So where
6 is—where is our city in regard to that?

7 DR. AMEN RA MASHARIKI: So one of the
8 things that we've done over the past year, which I
9 think actually was the catalyst for people hiring,
10 agencies hiring actual open data liaisons was that we
11 moved from sort of an every so often meeting with ODC
12 to a monthly meeting with ODC, and even in some
13 instances like I have alluded to, the agencies that
14 we know have a tough road to hoe, we sort of engage
15 them one-on-one. And so we had a very sort of
16 flexible scheduled engagement strategy with the
17 agencies where they're seeing that ODC is actually—
18 this is more of a full-time commitment, and there's a
19 lot more that needs to be done here. And so, we've
20 seen that move to sort of hiring of people. And I
21 think as we move forward that's the strategy that we
22 want to continue to grow within the agency is
23 bringing on people who are full-time committed to
24 overseeing. What you also saw was ODCs who were
25 there, over the previous couple of years, move on in

2 their job commitments whether it's to another agency
3 or somewhere else within that agency, and the people
4 that replace them, they're meant to be full-time. So
5 you're seeing that shift happen in many different
6 ways, but we're pushing aggressively for the agencies
7 to commit full-time ODCs because we're engaging them
8 in a full-time manner.

9 CHAIRPERSON VACCA: Does the PD have an
10 open data person?

11 ALBERT WEBBER: They do, yes.

12 DR. AMEN RA MASHARIKI: They do. The one
13 and there's a difference. Okay, and what do you do
14 when you note an error in one of the steps that you
15 posted? Have you noted errors? Has the public
16 brought your attention to errors? How do you find
17 out the errors and what do you do?

18 ALBERT WEBBER: So we've received—we—we,
19 you know, come to—come across errors in a variety of
20 fashions. One through comments in the Open Data
21 Portal. Through social media, sometimes people
22 Twitter accounts reviewed and they Tweeted these
23 errors. So I think it also depends on the type of
24 error. So I think errors that an agency may be aware
25 of, or that happens somehow through the automated

2 workflow, if we see that on our end, we can—we
3 certainly control it, and we'll work to correct those
4 errors. And I think there is also the—the fact that
5 errors that—that may come in as a part of operations
6 from an agency. So for instance a ticket for—is
7 accidentally with the data via 3016 as opposed to
8 2016. On our end we're not going to mandate an
9 agency changes their operations. So, if the agency
10 does find that error and points it out to us, and
11 wants it corrected, we'll work with them to put that—
12 that fix in place. But we won't mandate agencies to
13 change their operations because it's possibly a
14 clerical error that happened to go back in.

15 CHAIRPERSON VACCA: So if a member of the
16 public finds an error, who do they call? Where would
17 they go if they find an error? Should they call your
18 office or-or text you or email or what's the process?

19 ALBERT WEBBER: A comment. As of right
20 now, it's a comment on that data set.

21 CHAIRPERSON VACCA: A comment to

22 ALBERT WEBBER: Daily no. I'd say
23 weekly.

24

25

2 CHAIRPERSON VACCA: Weekly. So if
3 someone reviews, then you can—you can make that
4 chance.

5 ALBERT WEBBER: [off mic] Yes, but I—what
6 I'd like to point out is that with the new staff we
7 have having on board, that would be a part of the
8 things in their purview, review the comments, and
9 reviewing data set nominations, reviewing the errors
10 and just overall looking to clean up data on the Open
11 Data Portal. So as of right now, with our current
12 staffing, we're not reading the comments on a daily
13 basis. With the new staff, we will be increasing the
14 frequency in which there will be more comments and
15 requests.

16 DR. AMEN RA MASHARIKI: Can I?

17 CHAIRPERSON VACCA: Yes.

18 DR. AMEN RA MASHARIKI: Can I also add
19 that we make ourselves available in many different
20 forums for engagement. You will always see MODA and
21 DOITT's staff at Packathon (sic), at Data Solutions
22 at data events in and around the city where we engage
23 the people who are actually using the data, and
24 almost invariably we engage in conversations around
25 why I'm using this data set here for challenges and

2 seeing that data set. So that's—that's one form. I
3 have alluded to this social media, and so we're
4 consistently sort of watching our social media, and
5 people do tend to engage us there. Then there's also
6 via the—via the portal, and another mechanism that we
7 have is from MODA's perspective we've had many
8 businesses and uses of data specifically around sort
9 of, you know, B data and so on and so forth reach out
10 to us and say, hey, we'd love to come in, and sit
11 down and talk to you about what we're doing, and how
12 we're using open data and some of the challenges
13 we're seeing. So we make our doors open to any
14 organization that reaches out and wants to have a
15 meeting, and—and—and sit down with us. So we have—we
16 have any number of levels of the folks to begin to
17 engage finally with the community.

18 CHAIRPERSON VACCA: I wanted to bring up
19 about FOIL. We did touch on FOIL before. We have
20 agencies that are not compliant with the FOIL—FOIL
21 regulations. Sanitation is one of them. Buildings
22 is another one. Only about 30 agencies provided
23 their data on FOIL for the Open Data Plan, and why—
24 why they're not complying with the FOIL requirements

2 on Open Data, and what are we doing to get them to
3 comply?

4 ALBERT WEBBER: [off mic] So responses to
5 FOIL, to those FOIL kind of metrics are ongoing.
6 Since the plan was published on July 15th, there have
7 been a number of updates to that. We will get (sic)
8 bills updates for our members on the Open Data
9 Dashboard. So that there's transparency and it is
10 reported. I think what's happened over the last year
11 is that there's been—there's been a lot required of
12 our agencies and ODCs and to your point, ODCs sort of
13 need help. I mean we're asking them to really get a
14 dictionary for responses and an unprecedented number
15 of comments and the physical feedback. And I think
16 that agencies are working their hardest to provide
17 those numbers. With—with regards to—in some cases
18 this changes the way that FOIL is recorders had to I
19 guess track or review the FOIL requests that were
20 coming. So we're now asking you to look at from the
21 perspective of Local Law 11 whether or not the data
22 they released is a public data set or just a plain
23 data set, and I think there are some changes that
24 they're going to have to adapt to and go through.
25 But I do think the numbers that are confusing

2 decrease as, you know, agencies get used to this
3 piece of legislation.

4 CHAIRPERSON VACCA: I—I hope that what
5 you're saying is the truth because I—I have to tell
6 you that we noted that the Department of
7 Transportation has indicated that 521 FOIL responses
8 contain public data not on the portal. How can so
9 many instances occur? Five hundred and twenty-one
10 FOIL responses are not on the portal. That seems
11 like a very big gap. Is this agency being
12 cooperative with you, or is this a question of
13 getting the—the knowledge out there as to how to do
14 this, or when or whatever?

15 ALBERT WEBBER: It usually has been
16 cooperative. I mean they are an agency that probably
17 receives one of the highest number of public data
18 requests is on emissions. Even prior to open data
19 being in existence, I guess officially in the City of
20 New York, they had their own open data portal. So
21 they were already releasing data on a regular basis.
22 So we have a number of data sets from DOT. They have
23 reported on FOIL, and we've brought this then a
24 responsive data set requests. So by the letter of
25 the law, they've—they've been working very hard to be

2 compliant, and maybe it's in the required data with
3 all this other data. (sic)

4 CHAIRPERSON VACCA: Well, this number
5 stands out. There's something not going right here.

6 ALBERT WEBBER: [off mic] And I think—and
7 I think one thing that's going to help is before a
8 legislation is passed, we've opened up the
9 conversation before the release of data within the
10 agency. So now we have open data coordinators
11 involved, and general counsel, FOIL officers, and
12 communications officers. So we've increased the
13 conversation with all of these pieces of legislation.
14 So I do think that we will see that we need further
15 compliance by dealing with the other agencies on the
16 portal.

17 CHAIRPERSON VACCA: Now the more 109 of
18 2015 is the Response Timeline Law.

19 ALBERT WEBBER: Uh-huh.

20 CHAIRPERSON VACCA: And that specifies
21 that requests from the public must be addressed in a
22 specific three-month period. An effort was made by
23 MODA to comply yet almost all the requests are still
24 being marked open even if they've been answered, and
25

2 some of them still have not been answered. It always
3 seems like the good stuff. (sic)

4 ALBERT WEBBER: [off mic] So we're
5 continuing our work with agencies to response to
6 those requests. I think there--there are similar
7 requesting back in 2011, and I think that's a--I guess
8 some feedback for all our agencies to get to, but the
9 agencies that continue to send in their responses, I
10 think at this point we're around 75% when it comes to
11 issuing final determinations, and agencies are still
12 continuing to issue those--those responses. What
13 we're doing on our end to ensure that agencies are
14 completely aware of what needs to be done by the law,
15 as Dr. Ra Mashariki had mentioned hold ODC calls on
16 conferences to ensure that they're aware of all the
17 piece of legislation. We've actually--we have gone to
18 agencies and sat with them in some cases to sort of
19 help review nominations or any other data sets that
20 up there to help--help them comply as much as
21 possible.

22 CHAIRPERSON VACCA: Let--let me say this,
23 I appreciate that all the agencies are working
24 together collaboratively, and I do know that things
25 have gotten better as compared to when we had the

2 original legislation in my-originally at the site
3 hearings. So I'm here for as Chair of this
4 Committee, but that's--this is all about--this is the
5 reason I sponsored Local Law 8 because I wanted a
6 compliance mechanism. There has to be accountability
7 at a certain point. Now, I understand that DOI
8 recently approved the examination and verification
9 process that will be used by MODA to examine agencies
10 for compliance with open data. Now, can you describe
11 the examination process--process that will be used
12 based on that review by--the review of the assessment
13 done by DOI?

14 DR. AMEN RA MASHARIKI: Sure. So, it's
15 a--sort of a three-pronged process that merely some of
16 what DOI has--has done and also some of what MODA and
17 DOITT have done in the past. So it starts off with a
18 survey. So we push out to the respective agencies a
19 survey that has--and I can share some sample questions
20 on the survey. But we push out a survey and the
21 agencies are now required to fill--fil that survey
22 out. And it really sort of walks them through a--a
23 litany of questions around datasets, and we go
24 through-- We--we constantly cull the comments, and
25 request that on--online. So we'll put that, you know,

2 have you responded to this? Where is this data set
3 and so and so forth. So the agency has a period of
4 time to respond to that survey. Once we have the
5 survey and we've called that survey, then we push it
6 up for public comment. So then you make the public
7 to give the community an opportunity to ask questions
8 around well, what about this data set? I actually
9 thought about this dataset and so on and so forth.
10 And so when we do that, we use that as the process to
11 go back and forth. So, now we take the public
12 feedback and we use that process to go back and forth
13 with the agencies. But ultimately where we land, we
14 ask the agency to certify, and so we give them a
15 document to certify these are the data sets that we
16 now through this examination and verification. So
17 this is sort of the verification part, but these are
18 the data sets that through your own sharing of
19 information. What we've seen through requests from
20 the portal and then from public feedback is data
21 sets—public data sets that need to be made public,
22 and then it's their job to then certify and verify
23 and sign off on that, and then that becomes public.
24 So that's the process. You know, I—I think—can I
25 also add to the workflow that we have in place with

2 respects to accountability. The way you—we looked at
3 this was not only we had to keep moving on releasing
4 data sets as per the original Open Data Law, but then
5 we had seven new laws that we needed agencies to move
6 fairly aggressively and quickly to comply with. And
7 so we put that on their place, and we created a
8 strategy in partnership with the ODC to do that. And
9 so where you see—and I've—we've had full compliance
10 and communication and—and cooperation with agencies
11 on geospatial standards and data dictionaries and
12 others. And so where you see agencies really moving
13 aggressively to be helpful, we sort of take all of
14 that into consideration. So in some instances with
15 some laws, we—we—we want to ensure that there is
16 movement. Some agencies are moving a lot more
17 expediently than others, but we want to ensure
18 there's movement. The next step is I reach out to
19 the ODCs, and quite frankly maybe some of the
20 leadership within the agency to say hey, you're
21 moving—you know, we—we—we would appreciate it if you
22 moved more aggressively on this particular law or
23 this particular law. What we've—what we've refrained
24 from doing is sort of picking out one law where we
25 see an agency. They may be moving aggressively and—

2 and helpfully with others, and they're moving slow
3 on-on these. We sort of remind them—we sort of
4 remind them and continuously engage them. I speak
5 with the team almost daily around are we getting
6 movement? Is there passion there, or are we getting
7 no response? We're not getting anything, right, and—
8 and always the answer is we're moving. They may not
9 even when we pass this. (sic) Once I reach out, then
10 the next step is if we don't get the movement that we
11 want in order to be compliant then we push this—a—a
12 list to Commissioner Rose, and then the Commissioner
13 has Commission—Commissioner communication. So right
14 now that's the work. We refrained from sort of
15 pulling the trigger on that, so to speak, because
16 there is consistent movement. Speed can be faster in
17 some instances with some agencies, but there is
18 movement, and we're—we're appreciative of that and
19 working with it.

20 CHAIRPERSON VACCA: Okay, we've been
21 joined by Council Member Greenfield. Any questions
22 for the panel from the council members present?
23 Okay. No questions. I want to thank you so much for
24 coming, and enlightening us, and it's something that

2 our committee will be monitoring further, and I know
3 we'll be in touch with you. So thank you.

4 DR. AMEN RA MASHARIKI: Thank you very
5 much.

6 CHAIRPERSON VACCA: Another panel would
7 like to be heard. John Kaehny, Reinvent Albany;
8 Noel Hidalgo, Made in New York City; Joel—Joel whose
9 handwriting is terrible. [laughter] That's a video?
10 Oh, Noel, is this yours? Wow, you should be—you
11 should be a doctor. Joel Trinidad. Yes, come up.
12 Can you make one more seat? We'll have four. If you
13 would like to testify, you need to fill out a slip so
14 if you'd like—you might as well testify, they're
15 here. Okay. Alright, thank you so much and I have
16 to swear them in, right?

17 LEGAL COUNSEL: No.

18 CHAIRPERSON VACCA: Oh, no?

19 LEGAL COUNSEL: You don't need to.

20 CHAIRPERSON VACCA: I don't swear you in.
21 I just swear at you. Okay, he took it. [background
22 ground comment] Noel, do you want to start off? Or
23 John, do you want to start? Okay.

24 JOHN KAEHNY: Hi, my name is John Kaehny.
25 I'm Executive Director of Reinvent Albany, and also

2 Co-Chair of the New York City Transparency Working
3 Group and I'm testifying today both on my behalf and
4 that of my co-chair Gene Russianoff of the
5 Transparency Working Group and NYPIRG and of GWG
6 member groups, Common Cause New York and Citizens
7 Union. So first off, just thank you, Chairman Vacca,
8 and the Technology Committee. Really, the most
9 important part of my testimony is right now, and
10 that's just emphasizing to you how incredibly
11 important it is that City Council stay actively and
12 energetically involved in pushing, cajoling and
13 cheerleading for open data in this city. The origins
14 of open data in New York are as a partnership between
15 the Mayor, agencies, Council and the public, and your
16 role is absolutely crucial, and without you being
17 here, open data will not succeed in this city. So
18 thank you. Thank you very much. A couple of quick
19 comments. Really New York City just so everyone
20 knows is looked at a leader in open data. We are
21 watched carefully by cities all over this country and
22 the entire world, and what we're doing here is very,
23 very important for people everywhere. And in a lot
24 of ways what we're doing now is entering a new phase
25 of open data, but the phase that we're entering into

1 is that of fixing problems and listening to the
2 public and making sure that the data is high quality
3 and grows. The first part of open data is the part
4 that we started with in 2012, which was publishing
5 that data, and we've made tremendous strides now. So
6 as you heard from DOITT and MODA roughly three-
7 fourths of all the data sets that were called in that
8 original 2012 Open Data Law are now published and
9 online. DOITT is making great process in automating
10 data sets in particular where I have to point out New
11 York City is the gold standard for automating data
12 sets, and that's real important because it means that
13 automatically updated data once fixed at the agency
14 is then fixed for the public, too. So it saves lots
15 of time and money, and I want to point that out. I
16 also want to point out that open data is starting to
17 become part of the everyday conversation in New York
18 City. When you go to a City Council hearing or a
19 community board meeting you hear about 311 data. You
20 hear about NYPD and DOT traffic and crime data, and
21 it's really becoming a tool that more and more people
22 are becoming familiar with. So open data and the
23 idea behind open data is working in New York, and
24 that's an important thing for us to emphasize.
25

2 Ultimately, going forward we think the really big
3 payoff from open data is one that doesn't get talked
4 about a whole lot, but it's going to be saving city
5 agencies tens of millions of dollars by making it
6 easier for agency employees to access their own
7 agency's data, and that of other agencies. Because
8 like us and the public agency people have to go
9 looking and often googling to find data from other
10 agencies. So we think one of the big payoffs going
11 forward is to get as much data made public as
12 possible both for the public and government. We're
13 very optimistic, and I'm happy to be here today. The
14 City Hall has made a huge push behind Mindy Tarlow,
15 the Director of Mayoral Operations and Ann Roest, the
16 Director of DOITT. Probably over the last four or
17 five months, they've been vastly accelerating the
18 tempo in Administration's effort and investment into
19 open data, and they've called on many open data
20 stakeholders, expert, non-expert advocates and other
21 to get their ideas and to address long-standing
22 problems with open data. So we think that over the
23 last six months, the Administration, prodded in part
24 by the seven very important amendments that Council
25 passed. It has really, really started upping their

2 game on open data, and I'm very happy to be able to
3 report that today. Because open data is hard. It's
4 about transformation, and transformation is not
5 always easy, and I want to thank you for the
6 questions you asked today. I think they were right
7 on target, and the target is how does the public talk
8 back to the city and tell them when a data set is
9 wrong or has a problem or something like that. And
10 that's what some of the open data mimics were about,
11 and that's what I think a lot of open data oversight
12 and efforts are going to be about going forward.
13 That is interactive open data, and that was a real
14 huge shift in how government works because never
15 before have we tried to create some process some way
16 so that the public can tell agencies how to fix their
17 information. We don't—we don't even do that really.
18 So this is a pretty new thing, and we expect a lot of
19 growing pains, but it's a super exciting time and—and
20 we appreciate you being here in the committee and
21 pushing that area. Along with thanking Mindy Tarlow
22 and Anne Roest, I want to thank the guys that just
23 testified, Amem Ra Mashariki from MODA and Albert
24 Webber from DOITT. They really are incredibly
25 impressive and hardworking civil servants and

2 they're—they're guys to work with, and we feel very
3 privileged that the city has put so much talent into
4 open data, and that's a testimony for us that—that
5 City Council. I mean, pardon me, that City Hall
6 really does care about this, and is trying to meet
7 all of the laws' mandates, and really can meet the
8 spirit of the law. I just want to highlight a couple
9 of things real quick for you that happened this year
10 that we think are pretty cool. One is A+ to DOITT
11 for automating 100 data sets. Super important, and
12 we hoop they—hope they can keep going this year and
13 do the same thing. We give the Administration an A
14 grade for publishing a lot of important data sets
15 this year, the city budget, the city record online,
16 NYPD's seven major felony crimes and a huge tack in
17 (sic) is the information data set. We want to give a
18 shout-out to the GLC (sic). You asked who the
19 laggards were. We'd say they're one of the leaders,
20 and early adopters of open data, and they're showing
21 how great it is. We're also very happy that DOITT
22 and MODA are now leading and responding all of the
23 comments and requests on the open data portal. That
24 was a—a huge issue for our groups the last two years,
25 and we've very pleased to see that they're used in

2 the Open Data Portal. We're also happy that the
3 Administration is attempting to comply with the Open
4 Data Law as much as they can, and doing things like
5 publishing their annual update on time, which is
6 great because agencies don't always do that kind of
7 thing, and they've done so much extensive analysis of
8 open data user-users, including reviewing all the
9 comments dating back to 2011, and looking at some of
10 the web analytics. This is very high quality good
11 stuff. They didn't even brag about it, but we'll
12 brag about it for them. Some of the ongoing issues
13 that your questions touched, but I'm going to hit
14 them again is that there still is no clear process
15 for fixing public complaints about data quality
16 problems, and there's a lot of them. This is going
17 to be something that my colleague and—and many of us
18 are going to address again and again. But Mr.
19 Mashariki gave the example of Ben Wellington from
20 IQuant New York getting the police to fix how they
21 deal with certain parking tickets. Very few people
22 can do that. Very few people can call up the press
23 office or call up MODA directly and get them to fix
24 that directly. So we need a public process for
25 dealing with the fixed data sets. A lot of the bills

2 that the Council passed are direct hits. They're
3 totally appropriate, and we want to thank you again
4 for that. But some of the issues that they deal with
5 non-standard agency data, and the need for data
6 dictionaries are really hard. And you asked if the
7 city was going to be able to comply with all the
8 mandates. They're not. We don't think they're going
9 to be able to comply with the data dictionary
10 requirement. Not because they not going to try, but
11 because this is a big change, and this is tough, and
12 nobody really knows how to do it, and this is— In
13 the area of data dictionaries, but probably New York
14 City is being watched by more other governments than
15 any other. So how do you get data dictionaries that
16 work for 1,600 data sets and over 100 plus for the
17 agency? A big, big challenge, and we're going to
18 work hard with the city Administration, but we want
19 to flag that because we don't think that they're
20 going to hit the target. We hope they do, but we
21 think this is crucial because without a plain
22 language data dictionary—the data dictionary is the
23 thing that explains what the problems with numbers
24 are. That's what it is in plain English. Nobody
25 knows what the data is, and right now there's a lot

2 of data sets that nobody knows what they are except
3 for the three or four people in the agency that
4 produced this doc. So really they're gibberish or
5 they're a foreign language that has not been
6 deciphered yet. So a data dictionary is the Rosetta
7 Stone that unlocks the data for the public and for
8 other agency users. So we're going to be really
9 paying attention to that as super important. We have
10 some recommendations that we list out in some detail,
11 but I wanted to call out a couple things to you about
12 high-high demand data sets that are not coming forth,
13 and you asked about NYPD data. CompStat 2.0, which
14 is the Police Department's online crime website is
15 not open data, and it's strange because it's a
16 website that shows the data, and matched the data,
17 but the data that underlies it, the information is
18 not available on the Open Data Portal or NYPD
19 website, and we think it should be, and we'd like to
20 see Council push NYPD to make that available. So
21 it's kind of an odd form of transparency where you—
22 you put in online, but you don't provide the
23 underlying data. Another data set that we know is
24 very high demand is the DOT Paving and Milling
25 Schedule, which is in a PDF format, and we think that

2 you should push them to make it into a machine
3 readable format that can be mapped and put on line.
4 So those are just some specific things about what
5 agencies can do. I have a couple just quick
6 recommendations for this committee, and we have
7 written recommendation for the Administration that I
8 will not speak to right now. One, we'd like to see
9 you amend Local Law 7 of 20116, which is the FOIL
10 bill, a very small change that would require agencies
11 to name the data sets that have been FOILED. So we
12 don't just know the number of data sets that were
13 FOILED, but we actually get to have a list of them.
14 So we can see. In the case you mentioned of DOT,
15 what were those data sets that people were asking for
16 and that they FOILED for? And then when we see the
17 name, it allows the public and the Council to ask
18 okay, why don't you put them on the Open Data Portal?
19 Because that's the idea behind that law. So it—it
20 does seem like it needs to be very slightly amended
21 to name the data set requested, and whether or not
22 those data sets are scheduled for publication on the
23 Open Data Portal. We'd also like the Council to
24 consider mandating through legislation the creating
25 of a public data issues trackers. Like the federal

2 government's data.gov issues, which is what the
3 Federal Open Data Portal uses to track publicly and
4 transparently all of the open data questions they get
5 for their gigantic Open Data Portal. And we think
6 that might be a model here, and that Council should
7 look at that and consider whether or that needs to be
8 mandated. And lastly for you, we want to—I'm
9 circling back to what I said at the outset, I really,
10 really thank you and encourage you to keep going with
11 oversight hearings, and also to meet informally with
12 the City's open data community as much as you have
13 time for and your staff has time for. We think it
14 really, really matters and it helps public advocacy
15 tremendously to have you asking questions, and to
16 being engaged and interested. Thank you.

17 CHAIRPERSON VACCA: Thank you and I
18 certainly am willing to meet with the group again.
19 We met with them previously. I thank you for your
20 work. I do want to note that you said some nice
21 things about the city and the Administration and TLC,
22 and I think sometimes we don't know nice things. So
23 I'm glad you said that because I know the
24 Administration has worked hard on this, and I know
25 that TLC has worked and I appreciate their help. But

2 we tend to always articulate the negative side. I
3 think you. I—I do want to say that Mr. Mashariki is
4 still here, and that is a unique feature of the
5 meeting by the whole, because usually--and I'm here
6 in the Council more than ten years now—the
7 Administration people that testify leave when they're
8 finished and don't hear from the public. So I want
9 to thank him for staying and listening, and I think
10 that speaks to his level of commitment and
11 engagement. I did want to give you an experience
12 where in my own case I used open data. I had a young
13 gentleman unfortunately who was killed in a traffic
14 accident on a major stretch in my district on East
15 Fremont Avenue, and I was then doing research and I
16 used open data, and I found out that within four
17 years there were 100 accidents on this stretch, and
18 100 accidents in four years. DOT had proposed a
19 Vision Zero road common plan, and the community board
20 voted no, and then a year and half had passed, and
21 they had taken no further action. So I called DOT
22 in. I made some tweaks to what I thought the plan
23 should involve, some tweaks, but that generally I
24 supported the Vision Zero plan, and it's now been
25 implemented. I had been criticized pillar to post

2 in—and on literally Facebook this, and that. I've
3 been called every name, things that I have been
4 called I've never been called before, but I don't
5 think I am, and— but it's going well. And I think—I
6 think open data provided that impetus for my office
7 to do that type of research. You know, I know
8 communities that always say we want more police, or
9 we want more traffic agents. They often just go to
10 meetings to say we want more and more and more, but
11 they don't have the documentation. They don't do the
12 research. Here it's at their fingertips. Here we
13 can do the research. I do question whether or not
14 many of the community boards and the community groups
15 have the knowledge about how to access open data, and
16 I think that that has to be an inherent part of our
17 city's effort, and that is to train the community
18 boards first. They are city agencies. They
19 advocate, but they're advocating with one hand behind
20 their backs unless they have the training. I
21 understand that many of them will not to be trained.
22 Many of them do not think they have the time, and
23 many of them do not think they will ever be able to
24 be trained. Well, I hate to leave them behind, but
25 if that's the case, we have to move the boards ahead

2 because they're in the main line of making sure
3 people get city services. And when you go to
4 agencies asking for personnel or asking for service
5 delivery changes, this open data and what we've done
6 here at the Council is really their friend. So, we
7 do need that type of training, and I've been saying
8 that for some time, and I hope that we can get down
9 to the bottom of it. There really should be an
10 appointee from the agencies, or appoint people--
11 appoint people at DOITT who are going out to the
12 community boards and engaging them, and reassuring
13 that this--that you don't have to be a computer genius
14 to find out how to use open data, and that they
15 should be reassured. And they can--once they find out
16 it can help them, I think they will use it. Okay.
17 Next.

18 NOEL HIDALGO: I can't believe you said
19 all those things in preparation to my testimony. I
20 wonder whether or not you're psychic. So,
21 Chairperson and council members it's a great honor to
22 be here. My name is Noel Hidalgo. I'm the Executive
23 Director at BetaNYC. In front of you, you have seven
24 pages of kind of our detailed review of the
25 activities that we've done over the last year. As

2 I've been told before, I shouldn't read every single
3 word. So I'm going to just summarize it in—in kind
4 of a plain narrative voice. First is a huge thank
5 you to not this Council, but also to the
6 Administration and the work that the Open Data Team,
7 not only MODA but also DOITT, for their tireless
8 efforts as well as other agencies that show up at
9 these different types of events who actually take
10 their expertise, take the time to demystify what they
11 know, and how these technology systems have been put
12 in place. Our organization, BETA NYC is really
13 focused on figuring out how to demystify this
14 government technology environment that we're in, and
15 to use technology, data and design to hold government
16 accountable and to really—to give everybody an
17 opportunity to express their voice. So speaking of
18 where we are with the city's Open Data Law, in the
19 last 12 months our community has grown 700 new
20 members, which is kind of a spike in where we've been
21 since 2009. We're not over 3,700 civic hackers who
22 are really ready to kind of understand this
23 environment and—and put their talents to use to help
24 our neighbors. In the last 12 months, we've hosted
25 four significant events, which is I think kind of

2 helped pull back the curtain on where we went to go.

3 It was the NYC School of Data. We partnered with the

4 New York City Parks Department on Trees Count Data

5 Jam. We hosted two conversations, one on affordable

6 housing data, and another one with NYC 311 where they

7 released the city's second largest data set, I think

8 about 35 million rows of call entry day. DOITT was

9 also there with records on this. So through these

10 events, we really see that the open data community is

11 at a particular maturity that it needs more

12 investment. The seven pieces of legislation that

13 were passed are taking us in the right direction, and

14 they have really strengthened the city's open data

15 practice. Dr. Mashariki used ecosystem. I will

16 interchangeably use ecosystem and practice, and we—we

17 were honored as, you know, advocates and community

18 members to partner with the city and different city

19 agencies to help do these prototypes, and we think

20 that that is the appropriate step. I mean we are

21 partners, and so we were included. So, you know, we

22 can say that, but we really think that that's

23 exemplary of other communities that should be done

24 in—with other agencies. We commend the Mayor's

25 Office of Data Analytics for engaging all communities

2 as much as possible to gather that feedback and to
3 incorporate the City's data—data users' voice and to
4 the future direction of where we're going. And we
5 absolutely agree with everything that our friends at
6 the Transparency Working Group, Reinvent Albany and
7 NYPIRG through their testimony. One of the things
8 that you brought up is figuring out how to make data
9 usable for community boards. We absolutely embrace
10 the nation—the Administration's notion of open data
11 for all, and for the last year we've partnered with
12 Manhattan Borough President Gale A. Brewer, the
13 city's Open Data Fairy Godmother on figuring out how
14 to bring community boards into the 21st Century. And
15 we spent a year kind of looking at their—oh, (a) we
16 got a whole bunch of CUNY students. We put them
17 through a boot camp, and so we're starting to
18 prototype this open data education kind of
19 curriculum, and in that process we did a gap analysis
20 on community boards. And within Manhattan alone we
21 discovered that community boards have a desire to use
22 open data, but don't always have the bandwidth in
23 regards to how fast their Internet is, the education,
24 the tools, or the capability to process all that
25 data. And frustratingly, you know, there is zero

2 best practices out there on how to teach the city's
3 open data to itself. And luckily we partnered with
4 Pratt Institute out of Brooklyn through their Savvy
5 Program, and we've outlined a framework that we were
6 able to (a) apply to 311 data, and then apply to the
7 Trees Count data program, and so we have a—a good
8 understanding of what needs to happen in-in the way
9 forward. But we need to actually go out there and
10 teach more community groups to—to use this data, and
11 we need to figure out how to use the data.

12 Therefore, including this feedback and improving our
13 curriculum, but sadly, we haven't been able to find
14 any financial support to go out and teach the city
15 actually its open data. In regards to kind of the
16 things that we're starting to see within the seven
17 pieces of legislation, the—the Data Dictionaries Law
18 is a great start, and I love that the Administration
19 has stated over and over and over again that the
20 seven pieces of legislation are the floor not the
21 ceiling. And we feel that every data dictionary
22 should contain many tutorials explaining how to best
23 explore the data, and embrace the data portal's
24 functions. You know, in that, we also hope that the
25 Administration starts exploring other tools that

2 better suit kind of the diversity of users that we
3 have. One of them is a bulk data users. Our current
4 open data tool doesn't allow for bulk—easy bulk data
5 downloading, and that's a frustration that we've
6 heard time and time again. But at the same time,
7 we're confronted with a situation where the tool that
8 we have is the best to use for the general public.
9 You can go to the city's Open Data Portal. You can—
10 if you have a fast enough internet connection, if you
11 have a large enough monitor, if you have a fast
12 enough computer, and you have struck a little bit of
13 luck, you can actually sit there and play around with
14 the city's data. And frankly, this is kind of the
15 best tool that's out there, but frustratingly it
16 doesn't meet all of the needs of everyone. And so,
17 over the next year, we're going to continue to
18 explore kind of these insights in these seven pieces
19 of legislation that have been passed, and work with
20 the Manhattan Borough President, and the Fund for the
21 City of New York to figure out how we can develop
22 more insights within the next step of open data.
23 Some of our—some of the insights that we've already
24 gathered within this year is that we hope to see that
25 there is a workflow, a concrete workflow that

2 actively engages the public around what are the
3 issues within a data set, and how do you understand a
4 data set. We're starting to call this a user
5 centered data release workflow. We helped pioneer
6 this with the Parks Department and MOD and MOTI at
7 the New York City Parks Data Jam. And for those of
8 you who were geeky, it's a user-centered design
9 process that really explores kind of like how do you
10 start to uncover the key parts of a data set, and how
11 to best use the data set. The other big insight that
12 we had from our events this year was the New York
13 City School of Data. We featured—over 372 tickets
14 were sold. We checked in over 260 people. We had 18
15 sessions, 40 presenters. Sixteen of those presenters
16 came from New York City government, and we had three
17 elected officials. Council Members, we hope to have
18 you this year in celebration, or next in celebration
19 of the City's Open Data Law. One of the key
20 components of understanding this ecosystem, and this
21 practice is how do we include parents? I'm not there
22 yet, but I know that many of my colleagues are
23 parents, and so when we think about open data for
24 all, one of the core components is making sure that
25 we have events that are inclusive in making sure that

2 we can bring everybody in. And so, we encourage the
3 City to explore events that are inclusive of parents,
4 and make sure that their needs are also taken care,
5 i.e., offer childcare at events. So in conclusion to
6 our testimony, we have a detailed number of insights.
7 We are seeing a handful of agencies that are really
8 exemplary of the—kind of the technology practice as
9 well as the open data practice. New York City 311,
10 New York City Parks, Department of City Planning,
11 DOITT's GIS Division and the Taxi and Limousine
12 Commission are—are really at the forefront of what we
13 see where we want the different agencies to go. We
14 think that they—that those teams continue to be
15 supported and given the resources to lead by example.
16 I know that some of these agencies or departments
17 within these different agencies don't all—aren't
18 always seen as cutting edge, and that's a bit of a
19 frustration because we really want them to succeed,
20 and they really need our support. For those agencies
21 that get it, there's a huge opportunity now that the
22 data quality is there to do more civic engagement
23 events, which will lead to building better data
24 guides and—and—and kind of tutorials. We think that
25 the City is at the point where-- Well, this is kind

2 of like a tug of war psychologically. We think that
3 agencies should be accountable for their own data
4 release and that shouldn't necessarily be ultimately
5 on the backs of MODA. So that's the direction that
6 we would like to go, and to get there we want the
7 City to be outlining more standards, coming up with
8 concrete protocols, developing open source tools that
9 will allow agencies who are at the forefront to host
10 their data to really share their data in a—in a more
11 inclusive way. And fundamentally, through that whole
12 process, develop a practice, and that practice is
13 something that we have outlined at the very end of
14 our testimony. We're two years away from the—kind of
15 the—the deadline of 2018 of the City's Open Data Law.
16 We're starting to wonder about what is the future
17 legacy. What is the infrastructure? Dr. Mashariki
18 explained that there are three different agencies
19 that all have different hands within different
20 elements of the City's Open Data Law. It's great
21 that they've been able to glue all of those resources
22 together, and to be a well functioning, well oiled
23 operation. It's great to see them getting more
24 resources. But fundamentally, wonder how does this
25 last into the next successive administrations and

2 Council leadership. And so, we would love to see a
3 dedicated unit that brings together all of these
4 different components from leadership and development
5 of standards, technology and tools, evangelism to
6 education and trainings, and really make sure that
7 New York City is the number one city in the world in
8 regards to open data. This testimony is all built
9 with the things that we've seen from around the
10 world, and we think that with just a little bit more
11 of our shoulder to the wheel, we will get and be far
12 beyond everyone else. Thank you.

13 CHAIRPERSON VACCA: Thank you, thank you.
14 Thank you for your advocacy. Next. [pause]

15 JOEL TRINIDAD: Good afternoon. Thank
16 you for making me speak at this Council hearing. I
17 am Joel Trinidad (sic). I'm not a golfer, but I like
18 that game. I represent Open data. I'm a product of
19 the New York City's Open Data system. After winning
20 the gaps in 2011, I quit my job and I been open data
21 since them. I'm now part of the open data family,
22 and I'm here to testify in position of the City
23 Council on the need to modernize open data
24 infrastructure in the city. New York City is a city
25 like no other. I won't use your valuable time to go

1 through a long list of reasons, but I believe we can
2 all agree that one final reason why it's so, is that
3 it's the global center of so many industries,
4 finance, media, fashion, information technology,
5 advertising to name a few. Underlying all these
6 industries is the fact that they are all knowledge
7 based services driven data intensive professions, the
8 technology primarily. And if New York City is to
9 become the epicenter of the emerging urban
10 informatics industry governing gov tech, civic tech
11 and smart cities, I further submit that the current
12 Open Data Portal needs to be redone to support this
13 nation's industry. For we believe that open data is
14 essential infrastructure, and as such, it needs to be
15 open source, and we have proved that from the intense
16 support innovation and experimentation, not just
17 depending on part of them. And permit two federations
18 so that not only city agencies can participate in
19 growing these knowledge portals. But third parties
20 like universities, non-governmental organizations and
21 private sector—and the private sector can effectively
22 collaborate with the city in creating a robust data
23 commons open to public/private partnerships. Already
24 New York City is home to Civic Call(sic), BetaNYC,

2 Big Apps and all these innovations. It's also home
3 to two of the leading open source data portal
4 providers in the world, New Civic (sic), which is now
5 part of gov delivery, in your schools (sic) and for
6 the FY recently by Open Gov, a direct part of Big
7 Apps. We help governments around the world at every
8 level from national governments, state governments
9 and local governments use modern open source data
10 portal platforms. In my mind, an open source
11 solution allows the city to build internal capacity
12 especially now that it's retiring staff to tap, you
13 know, the genius of the Cloud, and the universities
14 could tack on in the cops-CUSP. Interestingly, NYU
15 CUSP actually stood up their C-Cam portal, and that's
16 great that they can do that. And it also accelerates
17 innovation and compliance to the recent amendments to
18 the City's Open Data Law. Again, the innovation, the
19 genius of the Cloud. There are specific things that
20 we can actually leverage. In Carnegie-Mellon
21 University they built a way to automatically compile
22 data dictionaries once you upload the data. In-in-in
23 Canada they built a way to do metadata creating,
24 which we're extending for the City of Boston. In
25 California, they have built a way for them to do data

2 set level permission at the data set level. So they
3 can use the data portal not just for public
4 transparency, but for sharing data with the workload
5 component within the agencies. And these are things
6 that are being built by governments, universities,
7 companies like ours. So this is a real ecosystem,
8 not just one vendor doing everything and you're
9 trapped. You got a roadmap of that vendor. And I'm
10 sure you're keenly aware of the urban innovation
11 leader, and an early adopter of technologies like
12 open data often comes the costs of creating Legacy
13 investments that quickly get leapfrogs especially so,
14 more so in the technology industry. So I submit that
15 the current solution is in dire need of a rehab, and
16 the City should consider stopping the Sole Source
17 Open Data Contract especially now that they removed
18 the New York City based providers in the same space.
19 Thank you for your attention.

20 JOHN KAEHNY: [off mic] I just wanted to
21 support—[on mic] Oh, pardon me. I just wanted to
22 support Joel's call for a review of the open data
23 platform, which is in our—our testimony. I didn't
24 get to it, but in the specifics. The Socrata (sic)
25 open data platform has some serious limitations and

2 it should be reassessed by the City's Open Data team
3 and—and counsel should inquire into what their
4 assessment would do.

5 NOEL HIDALGO: But fundamentally out of
6 all the tools we've seen the certain usability of the
7 Socrata data set—I don't want to stand—stand up here
8 and be the poster boy or poster child for Socrata.
9 There are some key features within the—the current
10 open data platform that enable the city to achieve
11 open data for all, and that is the—the fact that you
12 can upload tabular data, and within a browser, I mean
13 there are some caveats. As I said, you have the fast
14 internet connection, you have a big browser. You've
15 got to have a little bit of luck on a good day, but
16 that—our current provider of the portal enables us to
17 gout and to train to—to sit somebody down. Even on
18 this particular laptop, I can teach anyone how to
19 look at 311 data and give the context of community
20 boards, and—and this to me there is not a—it's not an
21 or situation. You've got to have one or the other,
22 and I think through our conversations that we've had
23 with the administration they understand that there
24 needs to be alternatives. There needs to be multiple
25 alternatives to servicing the city's open data, and

2 we BetaNYC we strongly encourage that exploration and
3 that investment, but not give up what we have now
4 because right now we get open data for all at least,
5 a bigger portion of the pie than if we were to drop
6 it and—and exclusively with some of the tools. But
7 we look forward and encourage an investment in
8 alternatives that are out there.

9 JOEL TRINIDAD: [off mic] This is an
10 important piece. (sic)

11 [background comments, pause]

12 TIO CHINO: I'm a member of New York City
13 Council Technology family. Councilman Vacca, thank
14 you. We are Tio Chino and I'm Tio Chino, and this
15 Daphne, and we are members of the newly created
16 Privacy Advocacy—Advocacy group related to privacy
17 especially related to online privacy. The majority
18 of our members are open data, open software, freedom
19 software, free software and transparency advocates
20 who advocate for the sharing of all New York City's
21 citizens data collected by public and private
22 enterprise, true government contracted programs like
23 LinkNYC or DOT. However, we don't believe privacy
24 should be an excuse for advocates—for agencies to
25 hide behind for not releasing open data.

2 DAPHNE: So we are here today to raise
3 aware ness of the problems that massive amounts of
4 public data collection can cause to the New York
5 City--to privacy of New York City citizens. We have
6 started an ad hoc advocacy group with the purpose of
7 creating an official New York privacy guidelines for
8 it. This official board could be composed of New
9 Yorkers from various fields including communication
10 technology, security and law experts as well as
11 citizens who would debate implications and make
12 recommendations to elected representatives when they
13 discuss technology contracts that include the
14 collection of private data, its safeguarding and
15 potential implications. Thank you for the
16 opportunity to speak before you today.

17 CHAIRPERSON VACCA: So you want to
18 protect our privacy. Tell me where do you see--where
19 do you see open data posing a danger?

20 DAPHNE: Well--

21 CHAIRPERSON VACCA: [interposing] Is
22 that--is that open data of your interest or your
23 concern or is it other aspects of technology?

24 TIO CHINO: It's--it's open data and
25 privacy in general. For example, we heard that the

2 TLC is the fore, leading of open data--data sets, but
3 the first data set included all the drivers' private
4 addresses, of all the drivers' license plus where
5 they live, which could lead to issues that we haven't
6 even thought. So, for example, Uber has sent
7 mailings to all the TLC drivers using those data
8 sets. Do we--do--should the city citizens who are, for
9 example, TLC licensees, be--their address be exposed
10 to the general public--

11 CHAIRPERSON VACCA: [interposing] I-I--

12 TIO CHINO: --around the world?

13 CHAIRPERSON VACCA: --I have to tell you
14 I did not that, and I-I do not agree that there are--
15 our private addresses should be posted.

16 DAPHNE: In Addition, with regard to
17 LinkNYC, if you look at the term--term of service,
18 it's very vague, and specifically with regards to
19 video, which supposedly right now is not active. But
20 when questions are asked, there are very vague
21 answers to some of these questions, and we would like
22 to get better answers on that, and also find out how
23 that would affected everyday citizens.

24 CHAIRPERSON VACCA: Well, if you have
25 concerns regarding LinkNYC, I'd appreciate you

2 sending me an email so I could have those itemized
3 concerns, and I would certainly ask the agency for an
4 explanation, and I do think that the TLC matter
5 should be looked into. I—I do not want people's
6 addresses-- I mean transparency is one thing, but I
7 do think that that is a stretch, and I don't—I'm not
8 happy with that.

9 TIO CHINO: Do they, the TLC—I—I looked
10 at the benefit of recent TLC data, and it's not
11 there, but in the past like a year or two year ago,
12 it was there.

13 CHAIRPERSON VACCA: [interposing] So it's
14 not the—all right let me qualify. It's not there
15 now?

16 TIO CHINO: Today's data set it's not
17 there, but the first data set they published was an
18 Excel spreadsheet, and everything was there.

19 CHAIRPERSON VACCA: Oh, then—then I'm
20 glad you clarified. It's not there now. It may have
21 been put there in error, and it may have been taken
22 down, and I do admit that with advent of open data,
23 there may have been mistakes, but I—I do want to make
24 clear that I do want to respect that—privacy rights
25 of individuals, and I—I will bring that up. We do

2 still have—Mr. Mashu (sp?) isn't here, and I'm hoping
3 that he will clarify that for the agencies that have
4 left, but certainly I want that to be noted. Okay.
5 I want to thank the panel, and there being no other—
6 Oh, we have one more? Oh, we have one more. Alisi
7 Billier—Alexis—

8 ALEXIS: [off mic] Alexis.

9 CHAIRPERSON VACCA: Alexis. I'm going to
10 write it today. Come on up. [pause]

11 ALEXA BILLIER: Hello, hello. Oh, that
12 is on. My name is Alexa Billier. I am not an
13 executive director of anything. I just—I am someone
14 who has a user story for share of the century. I'm a
15 CUNY Baruch College student, a mathematic major and
16 an enormous open data nerd, and this is a really—
17 really recent passion for me. I remember when I—way
18 back in February when I was just hanging out on line,
19 and I happened upon the—to see the page for New York
20 City School of Open Data. They talked about open
21 data, and the civic—civic—civic hacking all this
22 stuff, and I was like sounds like really cool. Like
23 how come I haven't heard of this before? And I went
24 to the—I went to this conference, and I was—I was
25 really hooked by the amazing work that BetaNYC is

2 doing, like Gale Brewer was doing, all—everyone—
3 everyone who was involved in this initiative was
4 doing, and I started to get interested in doing this
5 as well. So what happened is I started a blog where
6 I would essentially just look at open data released
7 by New York City, and I'd see like what are some
8 things I can find in it? What are some stimulations
9 that I can do, and big people would be interested in
10 looking at this data, and that effort has just been
11 initiative that has proven to be very successful.
12 It's something that's been modeled by work—on work
13 that's been done by other civic hackers like Ben
14 Wellington and civic technologists like Ben
15 Wellington and Chris Long. The website gets like
16 100—100,000 hits a month, and—and, you know, it's the
17 music from New York (sic) and the Polish and
18 Gothamists, you know, very popular music. So all—all
19 that is—is really great, and I want to say that, you
20 know, my—my personal story would not have been
21 possible without the work of the City Legislature.
22 It would not have been possible without the work of
23 the City—City technologists. It would not have been
24 possible without the work of—all the hard work that
25 has been done by everyone that's been involved in

2 this initiative, and they have been pushing this
3 conversation forward, and has really made New York
4 City a leader of all of open data, and for that I
5 think I—I really want to—before I say anything else,
6 I really want to thank everyone involved in that and
7 the City Council specifically. You know, I think
8 that as I look back on some of the things that we've
9 done, I think I kind of—I—I get excited about like
10 where we can go from here. Plenty of team (sic), the
11 Open Data Legislation were both taken in full force,
12 and I think in that—by that time we'll be seeing some
13 of these—these—these really larger really amazing
14 data sets that right now are siloed within
15 individual—siloed within agencies that if they could
16 release they would have just so much value. And I
17 know there are dozens of—of databases like these, but
18 for the purpose of the illustration I'll pick just
19 one, which is one of my favorite ones, which Biz Web.
20 The Department of Buildings has—has this application
21 for Biz Web where anyone can go on the—anyone can go
22 online, and they can look up any information they
23 want to know about a building that's interesting to
24 them. There are building some other building what—
25 whatever, and you can find out things like you can

2 find out permit status, vacancies, construction
3 details, addressing, violations, signage, elevator
4 records, plumbing, façade, boilers, cranes,
5 everything—everything there is to know about a
6 physical building is there. The trouble with it is
7 that this a—it's like a system that needs resources.
8 It was built in a different time. It was an
9 application that was built in 2001, and it's really
10 showing now. It doesn't run very—very quickly. It
11 has issues. It's built on top of a mainframe, and it
12 would be amazing if this data could be liberated, and
13 make public, made available to everyone in an easily
14 accessible and consumer format by open data. Because
15 that would mean that we can now also start with the
16 individual fillers. We can go through this data and
17 do some really amazing things with it. Biz Web
18 violations data could be used to find patterns of
19 abuse in New York City building management allowing
20 us to, for instance, practically model and send
21 inspectors to—to areas before disaster strikes.
22 Local legislators could use it to get weekly updates
23 of complaints in their districts, and allowing them
24 to get a better understanding of constituents' needs
25 before the phones start ringing. You know, civic

2 activists could build a modern information dashboard
3 so that anyone can go online and look up their
4 building and like find out interesting information, a
5 wealth of interesting information about the building
6 that they live in, or some other buildings that are
7 interesting. Private developers could use Biz Web.
8 They could lock into data to get a block-by-block
9 building-by-building history of construction in New
10 York City, and to help try and figure out where it's
11 going to go next. Construction agencies could more
12 into-more easily interface with this data to track
13 the permitting process for building, hoping to
14 accelerate the bill times. And finally, other
15 agencies that don't have easy access to this data
16 like for instance the Department of Sanitation or the
17 Office of Environmental Remediation who are
18 interested in using these data streams could finally
19 be able to go in there and use this data to
20 investigate problems that they face themselves, and
21 to streamline their own processes and even possibly
22 their own-their own legislative initiatives in a way
23 that-in ways that were never possible before. And
24 these is all, there are possibilities that are
25 unlocked by this one, you know, one-one example, this

2 Biz Web agent. There are dozens of applications like
3 this. Like for instance CompStat 2.0, which was
4 mentioned by John Kaehny, which are—have a wealth of—
5 a wealth of knowledge or wealth of opportunity within
6 them, and the only thing—the thing that we need to
7 to—to like use that is like make it public, make it
8 available, and a lot of—a lot of people use that.
9 And so what I think is most important how is—I think
10 that, you know, City—City Council has been the—has
11 been the leader in this field. The City council has
12 pushed—pushed the open data legislations, pushed
13 these—pushed these things that make this possibility
14 possible, and now it's up to us as a community,
15 government technologists, private developers,
16 legislators, everyone—everyone that's involved in the
17 Open Data Initiative to help push this conversation
18 further and to, you know, get more research that's
19 available to bring online these huge data sets. To
20 Bring online more—more—more information, more ideas
21 to do the hackathons, to do the idea sessions, to do—
22 to get the—get the juices flowing that help make open
23 data conversations that have made New York City a
24 leader in open data so far, and that will become ever
25 more important as we get closer to that critical

2 deadline when in-December 31, 2018 when-when
3 technically speaking all readily available data sets
4 should be available in the portal. And so these are
5 proven point things. And I think that I really want
6 to thank the City Council for the amazing work that's
7 done so far, and I really want you guys to continue
8 doing because this is-this is really important and
9 really-really essential. It's a huge opportunity for
10 social communication. I think it's a logic
11 opportunity. And 2018 is fast approaching. We will
12 face this pledge to do no less than to bring an
13 entire city's worth of resources online, and let's
14 make it happen. Thank you.

15 CHAIRPERSON VACCA: I thank you. It's
16 what you really-you really hit something on the head.
17 Let me say a couple of things. I use the Biz system
18 almost everyday, and I've been using it since 2003.
19 It's even before because I was a district manager to
20 a community board before I came here. I have found
21 Biz to be fantastic. I can check on demolition
22 permits, alteration permits, new permits, stop work
23 orders. So to me it's great, but I do think it has
24 to be updated, and many of the purposes that you
25 spoke about should be included in a revised Biz

2 system. Now, I do understand that the Buildings
3 Department is looking at that. I'm not part of those
4 discussions. However, but I do know that they're
5 looking at redesigning the Biz System, and as someone
6 who believes that this has been an--an unbelievable
7 asset, could you send me an email with your
8 suggestions, and then I have some--some of my own, and
9 I'd like to put together a position for the
10 commissioners for review based on my experiences.

11 ALEXA BILLIER: Absolutely

12 CHAIRPERSON VACCA: Isn't this great? I
13 think it's been belated. It's--it's slow.

14 ALEXA BILLIER: Yeah, that is.

15 CHAIRPERSON VACCA: I know, I know, I
16 know.

17 ALEXA BILLIER: If you want--if you want
18 information on these--

19 CHAIRPERSON VACCA: [interposing] The
20 government is slow, too. So, therefore, I'm kind of
21 used to this stuff, but otherwise--

22 ALEXA BILLIER: [interposing] Yeah.

23 CHAIRPERSON VACCA: --it's--
24
25

2 ALEXA BILLIER: [interposing] If you want
3 information on an individual building, this web is
4 amazing--

5 CHAIRPERSON VACCA: [interposing] Yes.

6 ALEXA BILLIER: --but it also, you know,
7 the interface is a little hard to get around, and
8 it's really hard to get information on--

9 CHAIRPERSON VACCA: [interposing] Well, I
10 have to tell you one thing I'm trying to do is I've
11 been trying to do almost since I'm here in the
12 Council, but that's something called a bad actor's
13 bill, and that is how do we trace bad actors who
14 don't maintain their buildings, yet they go and seek
15 permits to build new buildings on top of the building
16 that they already have, but they don't maintain, and
17 there is not that ability to cross-section that. And
18 that's something that I'm looking to do because why
19 someone be given new permits to build more when they
20 cannot maintain their property they have based on the
21 amounts of violations they have be they through HPD
22 or through--will though the Buildings Department.
23 They should be told no new permits until you
24 straighten out your act. Get the violations taken
25 care of, pay the city the fines the City is owed,

2 and then we will consider a new permit for a new
3 building. Right now, we don't have that capacity
4 across agencies, and I know you alluded to that in
5 your reform suggestion. So I'd appreciate something
6 to let me try to be helpful.

7 ALEXA BILLIER: I've got that in here.

8 CHAIRPERSON VACCA: Thank you. Thank you
9 very much. Okay, there are no further speakers. It
10 is now 2:35 and I want to thank everyone for
11 attending hearing of the Technology Committee. This
12 hearing is adjourned. [gavel]

13

14

15

16

17

18

19

20

21

22

23

24

25

C E R T I F I C A T E

World Wide Dictation certifies that the foregoing transcript is a true and accurate record of the proceedings. We further certify that there is no relation to any of the parties to this action by blood or marriage, and that there is interest in the outcome of this matter.



Date September 30, 2016