CITY COUNCIL
CITY OF NEW YORK

------------------------ X

TRANSCRIPT OF THE MINUTES

            Of the

COMMITTEE ON TECHNOLOGY

------------------------ X

                    September 20, 2017
                    Start: 1:50 p.m.
                    Recess: 2:29 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)

James Perazzo, Deputy Director Strategic Management
Mayor's Office of Operations
Acting Director of the Mayor's Office of Data
Analytics, a/k/a MODA

Albert Webber, Director of Open Data
Department of Information Technology
Telecommunications, DOITT

Mary Tobin Director
Brownsville Partnership

John Kaehny, Executive Director
NYC Transparency Working Group
Reinvent Albany

Noel Hidalgo

Sumata Hari Hurashura

2   [sound check, pause] [background

3   comments]

4   CHAIRPERSON VACCA: Okay. Okay, than you

5   everyone. Attention. Good afternoon. My name is

6   James Vacca, and I'm Chair of the New York City

7   Council Committee on Technology, and I'd like welcome

8   all of you to today's hearing. We are here to

9   discuss the Open Data Plan of 2017 in addition to two

10 new bills aimed at strengthening the Open Data Law,

11 and ensuring its continuation. [gavel] New York

12 City offers an immense number of services and

13 resources to residents, and when people take

14 advantage of these services, they are inevitably

15 involved in the creation of data. This data is

16 fundamental to the city's operations, and it's also

17 important that people have access to the data. Access

18 to government information ensured public institutions

19 can be held accountable. In New York City the

20 importance of public access to data is reflected in

21 both law and practice. In 2012, New York City became

22 the first municipality in the country to mandate that

23 all non-confidential government data be made

24 available online in machine-readable formats. Since

25 then, the Council has passed seven amendments

2 strengthening and updating this law, and the Open

3 Data Team had worked diligently to ensure citywide

4 compliance with the law. Under the Open Data Law,

5 the Department of Information Technology and

6 Telecommunications is required to public an annual

7 Open Data Compliance Plan, and for the previous three

8 years I have convened oversight hearing on the plan.

9 Today, is a little bit bittersweet as it will be my

10 final time chairing an oversight hearing on Open

11 Data. I must say in the time since holding my first

12 hearing I've been thrilled by the city's overall

13 progress. The city has consistently ranked at or

14 near the top of nationwide Open Data rankings, which

15 is not only because of the legislation the Council

16 has advanced, but also because of the work done by

17 the Mayor's Office of Data Analytics, DOITT, numerous

18 good government advocates, and the city's vibrant

19 civic tech community. This year's Open Data Plan

20 demonstrates the city has continued improving the

21 Open Data Portal, and made significant steps toward

22 increased public engagement. In the past year the

23 Council has created a new and improved Open Data

24 website for which they leveraged cutting edge human

25 centered design methods. The Open Portal has seen

2 170,000 new data sets added, 38 data sets automated,

3 and has engaged an average or 140,000 users per

4 month. The Open Data Team has also held number

5 well intended community events. Still, while there

6 has been tremendous success, maintenance of the Open

7 Data Portal and coordination between every city

8 agency is by its very nature a massive undertaking.

9 Issues are bound to arise, some of which we will

10 discuss today. Additionally, as technology advances

11 and agencies become more experienced with the

12 intricacies of data publication, there are several

13 ways we can improve current law. Over the past year

14 my office along with Manhattan Borough President Gale

15 Brewer, convened a number of stakeholders to solicit

16 feedback and Introduction 1707 resulted from these

17 conversations. A port of the new legislation

18 reflects currently—current agency practice, encoding

19 it into law and preventing any future administrations

20 from making changes. The Open Data Law currently

21 states that agencies should publish data sets before

22 December 31, 2018, and if data cannot be made public,

23 an explanation should be provided. Intro 1707 will

24 extend this deadline to 2021, thereby offering

25 agencies more time to comply. Another provision in

2   Intro 1707 will change the timeline for the

3   compliance plan pushing it from July to September.

4   The purpose of this is to better accommodate the

5   internal processes of agencies, and the Open Data

6   Team. This legislation will also require that the

7   head of each agency designates an employee to serve

8   as the Open Data Coordinator responsible for ensuring

9   that such an agency complies with the Open Data Law,

10 and receives feedback from the public. This will

11 codify a current provision in the Current Technical

12 Standards Manual. Additionally, this legislation

13 mandates that the department review the Technical

14 Standards Manual every two years, and provide a

15 method for public comment. Lastly, the legislation

16 requires DOITT to collect and publish data on the web

17 portal, codifying a current departmental practice.

18   In addition to 1707, we will be—we will

19 be discussing another piece of legislation, which I

20 introduced, Intro 1528. Currently, agencies must

21 review Freedom of Information Law responses, which

22 include the release of data, and determine if the

23 response consist of data that has not already been

24 included on the Open Data Portal, and to report on

25 these statistics of these responses to the Compliance

2   Plan—in the compliance plan. 1528 will require that

3   the Compliance Plan also includes the names of any

4   data sets provided in response t FOIL requests that

5   are not included on the Open Data Portal. The two

6   bill we're hearing today are a priority for this

7   committee, and I'm looking forward to moving these

8   forward in the next few months.

9   I want to thank our Committee Counsel

10 Maleta Jubali (sp?), and Policy Analyst Patrick

11 Moville for their hard work, as always in serving

12 this committee, and I want to thank my Legislative

13 Director Zach Hecht as well. I—I was going to call

14 upon the Manhattan Borough President to speak first,

15 but I don't' think she's here yet. We're joined by

16 members of the committee Council Member Barry

17 Grodenchik to my right, Councilman Joseph Borelli,

18 who's really to my right. Council Member Annabel

19 Palma. [background comments, laugher] Council

20 Member Annabel Palma as well. Okay, so why don't we

21 call the Administration? James Perazzo of MODA;

22 Albert Bueno of DOITT--[background comments] Webber.

23 I'm sorry. Your handwriting, that's okay.

24 [laughter] I shouldn't talk. It's okay. I have to

25 swear. I won't be swearing at you. I'll be swearing

2 you in. Please raise your right hand. Do you affirm

3 to tell the truth, whole truth, and nothing but the

4 truth in your testimony before this committee and to

5 respond honestly to Council Member questions?

6   JAMES PERAZZO: I do.

7   ALBERT WEBBER: I do

8   CHAIRPERSON VACCA: Thank you. Who would

9 like to proceed? Please introduce yourself for the

10 record.

11   JAMES PERAZZO: [off mic] I'm James

12 Perazzo, the--

13   FEMALE SPEAKER: Your mic isn't on.

14   JAMES PERAZZO: James Perazzo. Good

15 afternoon. My name is Perazzo, Deputy Director for

16 Strategic Management at the Mayor's Office of

17 Operations, and Acting Director of the Mayor's Office

18 of Data Analytics also know as MODA. I'm joined by

19 Albert Webber, Director of Open Data at the

20 Department of Information Technology and

21 Telecommunications. We are here today to speak about

22 the way MODA and DOITT have worked together over the

23 last year to fulfill the requirements of the city's

24 Open Data Law and implement the administration's Open

25 Data for All vision. On behalf of the

2 Administration, I'd like to extend gratitude to this

3 committee on the leadership of Chair Vacca for the

4 attention and care shown to—to open data over the

5 past several years. I'd like to begin by describing

6 the structure of the city's Open Data program. MODA

7 is the city's hub for advanced data analytics and

8 advocates for the use of open data in citywide

9 projects and in communities. DOITT manages the

10 technical operations of data publishing with [coughs]

11 with city agencies and our vendor partner Socrata.

12 Over 90 city agencies have a designated open data

13 coordinator who serves as the primary point of

14 contact with MODA and DOITT, and works with our

15 agencies' [cough] data owners to publish eligible

16 data sets. These three entities along with our

17 partners in the City Council and the civic tech

18 community work together to make city data open and

19 accessible to New Yorkers. (coughs) [pause] Since,

20 (coughs) the announcement of Open Data for All in

21 July 2015, the Administration has been unrelenting in

22 its efforts to put data in the hands of more New

23 Yorkers. We strive to make the Open Data Portal a

24 more user-friendly and accessible tool for anyone

25 interested in using city data consider a few

2   highlights since last year's hearing. The Open Data

3   Portal received over three million hits if Fiscal

4   Year 2017 with approximately 140,000 average monthly

5   users. The Open Data Portal now has a new website

6   designed in partnership with the open data community

7   through a human (coughs) centered design process, and

8   Chair Vacca pointed out. Thank you. The Park's

9   (coughs) Resource Centers piloted a data literacy

10   effort to teach New Yorkers how to analyze and map

11   the city's street tree census data. We celebrated

12   the five-year anniversary of the original pioneering

13   NYC Open Data Law in March with the city's first Open

14   Data Week produced collaboratively with the civic—

15   civic tech community. In partnership with the city's

16   Commission on Human Rights, MODA used several data

17   sets from the Open Data Portal to optimize where

18   inspectors search for landlords who are illegally

19   discriminating against potential tenants. In

20   addition, dozens of agencies continue to release high

21   value data sets, bring the total number of data sets

22   on the Open Data Portal to more than 1,700. We

23   published EMS Incident Dispatch Numbers from the Fire

24   Department, and detailed NYPD Complaint Data. We

25   also published the City Council's data set on

2 participatory budgeting, which is an open community

3 decision-making process that allow constituents to

4 choose how to spend public dollars. Each record in

5 this data set represents a project a New Yorker

6 nominated and got funded through the PV process. The

7 data set includes a plain language data dictionary,

8 and it's easily mapped against other data sets

9 through its standardized geospatial coordinates and

10 political districts fully complying with the 2015

11 amendments to the Open Data Law. MODA is currently

12 partnering with PV NYC to give this year's cohort of

13 participants, educational resources on NYC Open Data,

14 helping them to (coughs) better understand how data

15 sets from, for example, NYCHA, NYPD, Department of

16 Education, and other agencies can give context and

17 add specificity to their proposals. We highlighted

18 this and other examples (coughs) of the fall open

19 data impact cycle in the 2017 *Open Data For All*

20 *Report,* with MODA and DOITT submitted to City Council

21 and the public in July. Readers were also invited to

22 submit their feedback on a digital open source

23 version of the report, which received more than

24 12,000 hits in the two months since it went live?

25 The highlight (coughs) of the report was the stories

2 of the land use advocates, local start-ups, community

3 boards, CUNY students and interfaith organizers, New

4 Yorkers from all five boroughs who are using open

5 data to improve their communities. Our approach to

6 the report reflects our belief that showing the faces

7 behind (coughs) city data is how we spur even greater

8 use. The report highlights green infrastructure

9 researchers who use their analysis of open data to

10 start a conversation with the New York City Parks

11 (coughs) and the Department of Environmental

12 Protection about how a new model could improve storm

13 water collection in the city. It describes an

14 automated dashboard of neighborhood indicators used

15 by community partners to combat joblessness in

16 Brownsville, Brooklyn. These stories and others in

17 the report make up just a handful of the 140,000 New

18 Yorkers who visit (coughs) the Open Data Portal each

19 month, about 60,000 of whom are new to the platform.

20 In order to ensure we are not only increasing the

21 number of users, but also optimizing their

22 experience, we have undertaken significant efforts in

23 the past year to better understand their needs. When

24 results from the graduate capstone MODA facilitated

25 with Columbia researchers indicated that new users

2 didn't always know how to get started with the open

3 data -with—with open data. MODA and DOITT worked

4 with high frequency community users to design for

5 their needs. This resulted in a new look for the

6 Open Data Portal—for the Open Data Portal this spring

7 as well as a new customer service application that

8 allows the team to better respond to user's requests

9 while collecting (coughs) for help and—and correct

10 data on inquiry types, a new process that is

11 resulting in continuous improvement to the Open Data

12 Platform. The existing Open Data Law requires that

13 data from all agencies, not only select high

14 performance be published in a way that is accessible,

15 understandable and usable for everyone. Ensuring the

16 city's compliance (coughs) with the Open Data laws

17 and polices is the foundation of strong and just Open

18 Data for All. For this reason, we support (coughs)

19 the goals of Introductions 1707 and 1528. In

20 particular, with the extension of the Open Data

21 Mandate will allow our program to accommodate new

22 data sets that are created by city agencies. In this

23 year's Open Data Compliance Plan, for example, city

24 agencies identified over 100 new data sets that had

25 not previously been disclosed. This was not due to

2 any delinquency on the part of the agency's Open Data

3 Coordinator, but because new data systems are

4 constantly coming on line, and data that may not have

5 been previously qualified as a public data set is now

6 eligible for publication. Extending the legal

7 mandate to publish data sets past December 2018 is a

8 key priority for the Administration. Formalizing the

9 role of Open Data Coordinators and disclosing the

10 names of data sets used to respond to FOIL requests

11 will further enhance our goal of forming communities

12 of practice around open data city agencies. For

13 example, while the Office of the Mayor had previously

14 been serviced by a single Open Data Coordinator.

15 MODA worked with the Mayor's Office of Operations

16 this spring to recruit Open Data Coordinators for all

17 40 of the small offices that make up the Office of

18 the Mayor. Additionally, this year the Open Data

19 Team required that Open Data Coordinators worked

20 directly with their agency records officers and

21 [coughs] and general counsels to assemble their data

22 set inventories and compliance reports. We found

23 this to be an effective way to ensure that data

24 publishers were interacting with their counsel, and

25 as a result, we received a 100% reporting rate from

2   eligible city agencies. From street furniture to

3   road quality, and from taxi trips to parking meters

4   to small businesses, New Yorkers interact with their

5   city government everyday. Because we have a good

6   Open Data Law, this information is available for free

7   online. The extension of the law would allow us to

8   continue to ensure that as the city become more data

9   driven, it also becomes more open and transparent.

10   We are excited to continue to partner with the City

11   Council to extend the law and optimize technical

12   standards to allow open data to continue to thrive

13   the changing technological environment. I would like

14   to thank the Committee on Technology for the

15   opportunity to testify today, and for its ongoing

16   support of open data. I'd also like to thank Beta

17   NYC, the Bureau Blank, Cardo and Reboot, and all of

18   our many community collaborators whose support makes

19   this work possible. At this point, I'd like to turn

20   it over to Albert Webber, who will discuss the

21   Administration's perspective on the proposed

22   legislation in more detail.

23   ALBERT WEBBER: Thank you, James and good

24   afternoon Chair Vacca, and members of the City

25   Council Committee on Technology. My name is Albert

2 Webber and I am the Director of Open Data for the

3 Department of Information Technology and

4 Telecommunications, also know as DOITT. On behalf of

5 the administration I'd like to extend gratitude to

6 this committee under the leadership of Chair Vacca

7 for the attention and care show to open data over the

8 past several years, and for the opportunity to

9 testify today. We all share a steadfast commitment

10 to open data as an instrument of transparency and a

11 bridge to connect New Yorkers with their city. This

12 is why DOITT works closely with MODA to make open

13 data as accessible as possible. This year's Annual

14 Open Data Report was issued on July 14, 2017, and

15 details some of the extraordinary steps we've taken

16 to connect New Yorkers with the data being published,

17 and support them as they use it tackle issues in

18 their neighborhoods. In fact, one of our important

19 user groups is the City Council, which is why we've

20 collaborated with you to offer several open data

21 trading with Council members and their staff over the

22 next few weeks. I'm happy to report that these

23 sustained efforts have yielded real results. As you

24 heard, we are expected unprecedented engagement with

25 our data sets, which now number of 1,700 including

2   170 added in the last year. We have also worked with

3   agencies to automate 38 data sets ensuring that they

4   are as up to date as possible with some automatically

5   updating as frequently as every five minutes such as

6   DOT's Real Time Traffic Data Set. The Open Data Team

7   continues to make great progress every year, but the

8   expiration of the Open Data Law, Local Law 11 of 2012

9   at the end of next year, gives us all the opportunity

10   to codify some helpful technical changes.

11   Introduction 1707 addresses many of these useful

12   fixes as a result of a productive and collaborative

13   effort between Chair Vacca, Borough President Brewer,

14   Beta NYC, Reinvent Albany, MODA, and DOITT. The

15   proposed bill would:

16   1.   Codify a bi-annual review of the

17   Technical Standards Manual, also know as the TSM. The

18   TSM is the dividing document on how agencies manage

19   and present their data. These standards are designed

20   to make open data more usable to the maximum number

21   of users. We see the TSM as a living document meant

22   to keep up with the advances in technology, data

23   availability and the passage of local laws that

24   impact data disclosure. An official periodic review

25   of this document will be sure that open data

2   disclosures stays up to date, relevant and true to

3   the most current practices.

4   2.   Extend the Open Data Law another

5   three years. Current law requires agencies to

6   publish and prioritize for release of their data sets

7   by December 31, 2018. We want to ensure that

8   agencies continue to schedule new releases beyond

9   that date. While simply extending this sense that

10 it's helpful, we would like to discuss how to improve

11 the language to ensure that submissions meant to be

12 disclosed by December 31, 2018 will still be held to

13 that deadline.

14   3.   Change the deadline for the Annual

15 Compliance Plan from July to September. The Open

16 Data Team has learned a lot about the process of

17 collecting information from Open Data Coordinators as

18 the July 15 deadline approaches. We have found that

19 this timing bumps up against budget season, and

20 request that it be changed to provide sufficient time

21 in each new fiscal year to gather and report

22 information. Additionally, the annual Mayor's

23 Management Report, MMR, is published in September.

24 Shifting the deadline to September will allow us to

25

2   align these two reporting exercises and efficiently

3   deliver a unified report representing our progress.

4   4.   Codify the requirement for each

5   agency to designate an Open Data Coordinator. While

6   Open Data Coordinators are currently required under

7   the Technical Standards Manual, codifying this

8   requirement would underscore agencies' commitment to

9   open data in perpetuity.

10   5.   Require web portal site analytics in

11   perpetuity. Since we already publish web analytics,

12   we support the idea of ensuring that future—future

13   administrations do the same.

14   In addition to these improvements, we

15   suggest amending the license provision of Local Law

16   11 of 2012 to best serve the goals of Open Data for

17   All. Currently Section 23-502 states that all public

18   data sets on the portal should be made available

19   without any license requirement or restriction on

20   their use other than attribution and description of

21   the modifications made to the data set. We have

22   interpreted this to mean that DOITT cannot make the

23   data available under any license including permissive

24   licenses. However, our mission to encourage public

25   engagement with open data in the long term makes the

2 ability to license very important. Structured

3 properly licenses could allow us to formally make the

4 city's data sets free and open for public use in

5 perpetuity, codifying our existing practice and

6 ensuring New Yorkers have access to open data for

7 long haul. We look forward to a productive dialogue

8 about this provision.

9   Introduction 1528 would amend Local Law 7

10 of 2016, which currently requires agencies to review

11 Freedom of Information Law, FOIL requests containing

12 data to determine whether they contained new public

13 data sets that could be published on open data.

14 Local Law 7 requires agencies to disclose metrics on

15 FOIL responses. Intro 1528 would require agencies to

16 take the extra step in gaining the data sets relating

17 to FOIL responses. This legislation furthers the

18 Administration's transparency objectives. Thank you

19 for the opportunity to testify today. Open data

20 remains a priority of this Administration shining a

21 bright light on our government and out city for all

22 New Yorkers to see, and providing the tools to solve

23 thorny civic problems in creative ways. We thank our

24 partners in the City Council and the civic technology

25 community for their continued advocacy. This

2   concludes our prepared testimony, and we look forward

3   to answering your question.

4   CHAIRPERSON VACCA: thank you very much.

5   I'd like to welcome Council Member David Greenfield

6   who has just joined us. Thank you for work that

7   you've done, your agency has done. We do think that

8   open data--open data is a transparency tool that's

9   fantastic, and we want it used more. We want to urge

10   people to use it on an increased basis. Let me ask

11   you some questions. What kinds of reasons prevent

12   the data sets from being automated? Is this because

13   of pushback from agencies, or the stress of techno—

14   technological limitation?

15   ALBERT WEBBER: It depends. So, it

16   depends on the frequency of update. So, there's a

17   lot of work that goes into setting up that automated

18   feed that, you know, delivers data automatically from

19   the agency's data base or from wherever they have it

20   stored to the Open Data Portal. So, we believe that

21   data should be—if it's updated monthly and more

22   frequently, those are things that we'd like to

23   prioritize for automation. So, if you have something

24   that's automated, say for instance on an annual

25   basis, the work that goes into setting up that

2   automation doesn't always make sense to deliver that

3   data automatically. What ends up happening in some

4   cases that we found out early, and this is something

5   that we've—we've tested, will have a data set that's

6   automated once a year. Over the course of 365 days,

7   there's a chance that, you know, that something could

8   come in the way, and the work that goes to setting

9   that up is not worth it. It's easier to just take a

10   data set, email it, have us load it as opposed to

11   going through a significant process just to set up

12   that automation.

13   CHAIRPERSON VACCA: Well, I do see that

14   38 data sets were automated this past year. So, how

15   does your team decide which to automate?

16   ALBERT WEBBER: It's a combination

17   between us, and some from agencies on it. Any agency

18   that wants to automate data will automate it. We

19   take into consideration the public comes back

20   heavily. So, there are certain data sets that we

21   know are getting heavy usage or if we're getting live

22   requests for them, we'll prioritize those for

23   automation. We'll automate anything that agencies

24   are will and able to automate with us.

25

2   CHAIRPERSON VACCA: Well, you did mention

3 public feedback, and I wanted to know how often the

4 comments on the data sets are checked. How often do

5 you look at the comments that the public makes

6 concerning the portal?

7   ALBERT WEBBER: Yeah, we're looking at

8 them everyday. So, this past March when we updated

9 the Open Data Portal, one of the things that we did

10 was implemented a new tool that takes in public

11 feedback. One thing that—that wasn't as available on

12 the site prior to the change was a direct way to make

13 feedback first specific use cases. So, now, you can

14 make comments on public requests, data question, data

15 errors. We found that there are a lot of questions

16 that came in from the Taxi and Limousine Commission.

17 So, we make it apparently that if you want to make a

18 request for a certain thing, it's there. The team is

19 monitoring these daily, and the five or six people

20 monitoring that everyday.

21   CHAIRPERSON VACCA: How many comments do

22 you get on a typical week?

23   ALBERT WEBBER: It varies. We—I could

24 probably get back to you those stats. We have it

25 broken down to different categories, and they're

2   monitored by different groups and, I—I want to say

3   roughly 30 or 40 a week, but it's hard to say. I

4   could get back to you with those numbers.

5   CHAIRPERSON VACCA: Okay. Let me ask you

6   when you deal with agencies concerning open data,

7   what's a typical issue that they raise? What—are

8   there agency concerns that we should know about that

9   we typically get when you talk to them, and try to

10   make sure the rule is implemented in its full, you

11   know--?

12   ALBERT WEBBER: At this point the

13   engagement is pretty smooth. The law was passed in

14   2012, but we've been doing open data since 2009. Me

15   and some agencies had their own sites prior to that.

16   So, at this point, it's--it's—it's fairly smooth of

17   an engagement. We try to provide as much documents

18   (sic) as possible for agencies in terms of the data

19   that they should be releasing. They're looking at

20   public feedback, and they're looking at their

21   websites. So, what—it's pretty smooth an engagement

22   at this point.

23   CHAIRPERSON VACCA: But we did a review

24   of the Open Data Portal for the past four years, and

25   it shows that some agencies are not reporting

2   properly. Who's overseeing the agencies to ensure

3   that they are in compliance?

4   ALBERT WEBBER: DOITT and MODA both

5   provide guidance to agencies to work with them to—to

6   work towards compliance.

7   CHAIRPERSON VACCA: Have you found

8   resistance, or is there an issue with this? What

9   people—are the agencies--?

10   ALBERT WEBBER: Yeah, they run into

11   resistance. I think we have the right Open Data

12   Coordinators in place, and the right people in every

13   agency that are making the right decisions or talking

14   to the right people. So, we have not met resistance.

15   CHAIRPERSON VACCA: Well, I realize when

16   you deal with agencies we're talking big, small,

17   medium size commissions and the gambit, but we did

18   look into the Civil Service Commission, for example.

19   So, the Civil Service Commission updated the plan

20   with their appeals data every year since 2013, except

21   for 2016. The 2016 data was neither included in the

22   portal, and nor was the removal of the data. We

23   caught it in the 2016 Data Report. So, how will MODA

24   ensure that these plans are being reviewed for

25   accuracy year-to-year?

2   ALBERT WEBBER: We can look into this

3   particular instance, and every year in the plan what

4   we include is an agency comment with the scheduled

5   releases. So, if the data has changed, if a data set

6   is—if a change in priority is made, we'll include

7   that comment in the plan, and the same thing with

8   removals. If a data set is removed from the plan,

9   we'll include the reason why, and if that was

10 removed, that—that would have been a mistake and

11 we'll definitely look into that.

12   CHAIRPERSON VACCA: So, I want to make

13 sure that we guarantee that even the smallest of

14 agencies are complying with the law. I know that

15 there are personnel issues possibly in the smaller

16 issues, but that can't be a subterfuge in—in my view.

17 Now, NODA—MODA indicated that they were working on a

18 survey at the last data hearing we held. So, that

19 all agencies can submit their public data to the

20 portal more—more systematically. Has the

21 Administration taken other steps to ensure that

22 agencies are in compliance?

23   JAMES PERAZZO: Well, I would say this is

24 part of the reason that we greatly expanded the

25 number of Open Data Coordinators, including all he

2 smaller offices of the Mayor and having those

3 specific data groupings

4   CHAIRPERSON VACCA: Okay. I wanted to get

5 to this now [mic squawking]. In this year's Open

6 Data Report, I was very happy to see that MODA will

7 be releasing an on-line project library, and

8 documentation of its Open Source Analytics framework.

9 As you know, I'm very concerned with city agencies

10 develop algorithms, and this is an important step in

11 the process. Do you feel this type of documentation

12 is something that could and should be encoded into

13 law?

14   JAMES PERAZZO: Well, I—I think we're

15 looking forward to a hearing on that topic--

16   JAMES PERAZZO: [interposing] I know.

17   CHAIRPERSON VACCA: --correct?

18 [background comments]

19   JAMES PERAZZO: I—I would have to refer

20 to that to speak at that—at that time but, of course,

21 MODA does believe that documentation is—is very

22 important is and communication is an important piece

23 of that Audex. (sic).

24   CHAIRPERSON VACCA: Okay. My concern is

25 transparency so that people can understand how

2 agencies arrive at the decision they arrive at, and

3 that's where al—that's where algorithms come in, and

4 I do know we're having a future hearing on my bill

5 next month, but as—as a preview it's kind of

6 inconsistent that we go with an open data outreach,

7 an open data engagement and then people still left in

8 the dark when it comes to other parts of government

9 decision making. So, I'd like you to look at your

10 Open Data Report and then get back to us because I

11 think that that report sets a framework for us to

12 have those serious discussions we should have on the

13 algorithms. So, we have to be consistent, and that's

14 why I brought that up. I know I can expect support

15 because people are entitled to that information. You

16 know, I come back to the same old story. So, I have—

17 I have many youngsters in my district, and they

18 applied for high school, and they may reject it.

19 They get their seventh choice, and they're told that

20 the school you want has no room. That's the

21 explanation. They don't know what went into that

22 decision. What information was fed into a computer

23 that didn't give this kid his first, second or third

24 choice, but gave somebody else their first, second or

25 third choice? It's basic transparency that parents

2   are entitled to, which they don't get right now.

3   Okay. We'll go back, but I wanted to cite that.

4   Now, are there any staffing budget needs that are

5   unmet by your agency that you feel should be

6   addressed by the Council or the Administration? I

7   know that we're doing a lot of work here, but I—do

8   you feel that the staffing is adequate based on the

9   outreach or the administrative work that's involved?

10   JAMES PERAZZO: At this time I—I think

11 we're meeting our operational requirements with the

12 existing staffing.

13   CHAIRPERSON VACCA: Okay. This year the

14 open data Help Desk feature was added was to the

15 website, and there was a large increase in the number

16 of inquiries made. Are the complaints made through

17 the Help Desk available to the public to review?

18 Will they know that the complaint they made has been

19 raised and addressed? How can we track that, track a

20 complaint?

21   ALBERT WEBBER: So, of the feedback that

22 comes in, we're currently making the public data set

23 requests available, and we also have an alerts

24 feature, and so when a complaint or a data error

25 comes in that is extremely of note, we'll make that

2 available. So, if for some reason a data set if the

3 automation went down, and it's down for a period of

4 time, longer than we'd like, we'll make those alerts

5 available.

6   CHAIRPERSON VACCA: Okay. Okay, I have

7 no further questions. [background comments] No? Oh,

8 yeah about the legislation. I forgot. [laughter]

9 So, sometimes. So, you seem to be in favor of—you

10 have no problems with my bill as is, the bill I

11 introduced?

12   JAMES PERAZZO: No.

13   CHAIRPERSON VACCA: You—your testimony

14 indicates that as is 17 what? 07. I forget the

15 numbers. On 1707 you have no problems with?

16   ALBERT WEBBER: We support the goal of

17 the legislation.

18   CHAIRPERSON VACCA: No, no, no, no, don't

19 start with that. [laughter] Don't start with that

20 with me. You said you had—you support the goal. I

21 mean we all support the goal of winning the lottery.

22 I doesn't mean we win. [laughter] Do you—do you

23 support the legislation? I got the impression.

24 Should I be reading between the lines? Are—are you

25 in politics that I have to read these things so

2 carefully because politicians are always, you know,

3 they never say what you think they said. But let me

4 ask you where are on that.

5   ALBERT WEBBER: We largely support the

6 legislation.

7   CHAIRPERSON VACCA: Do—do you have any

8 amendments that you would like or anything that

9 concerns you because that's what these hearings are

10 about also the leg.

11   ALBERT WEBBER: We mentioned that these

12 in terms of the licensing provision on the Open Data

13 Law. We—we do feel that changes to that legislation

14 where we could potentially include license

15 provisioning on the Open Data Portal would be

16 helpful. The way it's currently interpreted by us is

17 there's no license that could be put on a data set,

18 and will be very helpful.

19   CHAIRPERSON VACCA: Okay. Now, since

20 introducing the legislation, we've heard from

21 advocates that the extension of the deadline to 2021

22 might have negative implications and that agencies

23 will potentially lack the impetus publish data sets

24 sooner. Do you have that—do you have that concern?

25

2   ALBERT WEBBER: No, my—no. I think

3 looking at—putting—always putting a deadline I think

4 is always helpful, but now we do not have concerns

5 that pushing it to 2021 will delay the release of

6 some data, sets and—and we would also like to look

7 at—looking at the legislation or looking at the terms

8 of the legislation to make sure any data that's

9 solicited between now and 2018 is still held to that

10 deadline.

11   CHAIRPERSON VACCA: Okay, that's

12 important.

13   ALBERT WEBBER: To us as well.

14   CHAIRPERSON VACCA: Now, on 1528 that you

15 that there are 15—5,400 sponsors that include public

16 data sets not included on the Open Data Portal this

17 past year. Would there be any technical issues with

18 releasing the names of these data sets?

19   ALBERT WEBBER: We don't believe there

20 will be.

21   CHAIRPERSON VACCA: Now, concerning

22 Council Member Brewer's bill, you—you would have some

23 request to-- I—I think you did say you were for that

24 in—in principle but you had some issues there.

25 [background comments] That's 1707. One was the

2 framed bill that was mine. The other bill was

3 Council Member Brewer's bill. [background comments]

4 I'm also on that. I know, but is it determining the-

5 -if you take which—which-- What was—what was the

6 number? [background comments] How is 1707?   You

7 said here that you would codify a bi-annual review.

8 [background comments]

9   ALBERT WEBBER: Yeah, that was in regards

10 to the licensing provision.

11   CHAIRPERSON VACCA: That's the licensing

12 provision?

13   ALBERT WEBBER: Yes.

14   CHAIRPERSON VACCA: I got it then,

15 alright, now regarding 1528, you have no problem with

16 that bill? I want to clarify with you on both

17 pieces.

18   ALBERT WEBBER: Oh, yes. No, we—we—we

19 are to support it.

20   CHAIRPERSON VACCA: Yeah, because you

21 spoke about licensing requirement. I—I didn't get it

22 from there because my bill does not have registrants,

23 and my bill is 1528. Under my bill will give you—you

24 have issues with it?

25   ALBERT WEBBER: So, in general--

2   CHAIRPERSON VACCA: [interposing] You

3 have some concerns regarding licensing?

4   ALBERT WEBBER: Yes, the only thing with

5 the licensing is that we'd like to see some language

6 that would allow licensing. The way it's being

7 interpreted is that we can't even have permissive

8 licenses on the data. So, if there could be some

9 changes where we could do that that would be

10 extremely helpful.

11   CHAIRPERSON VACCA: Okay, I got it.

12 Okay. There being no questions from the Council, I

13 will thank you so much for coming, attending the

14 meeting, and thank you for your help. I've got to

15 say this is our last hearing of this committee, and I

16 hope the next committee grills—grills you like have.

17 [laughter] So, let's put it that way. Okay, thank

18 you so much.

19   ALBERT WEBBER: Thank you.

20   JAMES PERAZZO: Thank you.

21   CHAIRPERSON VACCA: Thank you. Did it

22 work? Is Gale here? Oh, Gale Brewer is here. Okay,

23 Borough President. I was going to say Council Member

24 but—

25

GALE BREWER: [off mic] You didn't tell me.

CHAIRPERSON VACCA: I did, right, Council Member—former Council Member Gale Brewer, the Manhattan Borough President is a prime sponsor of one of the bills today, and she—her work—her work as a pioneer of open data is well known to all. So, we welcome you as always.

GALE BREWER: Thank you very much. It is a pleasure to be here. I am Gale Brewer. I'm the Manhattan Borough President, and I want to thank Chair Vacca, and—for all of his support on this topic. I want to congratulate him in the New York Times article talking algorithms. That was very exciting. I'm not the expert. There are experts in the room, but I do think that our local law 11 of 2012 the Open Data Law was a turning point in some of these discussions, and you should know, Mr. Chair that I just met with a—elected officials in Spain, Madrid. Maybe others have met with them. I know the city did and they are looking at our law. They passed an ordinance, and we can talk about the reasons we have we have laws instead of ordinances, but then this gentleman was going to be speaking at

2 the United Nations yesterday about the issue of open

3 data. So, it's a very current topic around the

4 world, and I think we should be proud, not just me,

5 but apparently our law is considered a model for

6 looking other cities. So, congratulations to you for

7 keeping the discussion going, and it is always

8 exciting to be with all the civic tech organizations

9 that are here. So, no question, and you said this

10 just now that whoever unfortunately will not be here

11 as chair, but the Open Data Law continues to be a

12 major undertaking and questions must continue to be

13 asked. I want to thank you the unit, and I want to

14 thank the Chief Technology Officer, Miguel Gamino for

15 all of their efforts and MODA, people you have

16 mentioned, and they're all helping keeping this law

17 current. It was Mayor Bloomberg who signed the law—

18 the bill into law, and that was, as I said, the first

19 large scale municipal open data portal. There wasn't

20 a regulation on the ordinance, and I think it's—it's

21 an important point to make because we know that as—

22 just after the inauguration of a certain federal

23 official on January 20, 2017, the federal government

24 began removing information from the its websites that

25 had been published in the name of public access and

shared knowledge. There's no federal law to protect
public information from an administration that does
not believe in facts or transparency. You know,
that—that's why we have a law. By contrast, we are
assembled here today to discuss making our
transparency laws stronger and more effective, thanks
to the chair. I certainly support your Intro 528,
reporting the names of public data sets that are
requested under the Freedom of Information Law is
within the spirit of FOIL and open data, and I
support amending 1528 to allow DOITT to consider
which license will make the city's data sets free and
open for public use in perpetuity, and on the –that
we both introduced (sic) 1707 to further integrate
open data reporting into workflows in culture of our
civil service while giving advocates of transparency
more tools to analyze and respond to the city's open
data methodology. 1707 would support the work of the
Mayor's Office with the Open Data Coordinators who
guide their agencies towards compliance by making the
position a required one for each agency. Along with
open data—data advocates likely Reinvent Albany and
Beta NYC both of whose leaders are here today, thank
goodness. The Open Data Coordinators are the unsung

2 heroes who make New York City leader in open data.

3 That's where we are today, and I want to congratulate

4 everyone. The new requirements in this bill for the

5 collection, analysis, and public reporting of Open

6 Data Portal analytics will support your efforts and

7 strengthen your legacy. All legislations need

8 tweaking to enhance potential, and obviously this

9 bill is no different. I know that the text creates

10 some confusion. We talked about this earlier about

11 the 2021 deadline for agency comply—compile—

12 compliance. I think the language could perhaps be

13 clarified to state the original 2018 compliance

14 deadline is still in effect, and that Intro 1707 is

15 extending the law itself to 2021. The Open Data

16 Coordinators of each agency are working very hard to

17 meet the 2018 deadline, and give New Yorkers a full

18 selection of what the city is able to publish. I

19 would like to see added to the bill a safeguard that

20 we have all worked on to build and improve where New

21 Yorkers to one day elect a mayor like the current

22 leader of our federal executive, open data would be

23 endangered. We know now that our democracy is only

24 as strong as the institutions that serve as its

25 pillars and the laws and advocacy support them. I

2 believe you must provide a right—a private right of

3 action to protect New York City municipal open data

4 apparatus from a future administration that does not

5 wish to operate it in good faith as the current and

6 previous administrations have done. I had talked

7 last year when I had talked about our then future

8 plans to work with Noel Hidalgo, who's here. He is

9 the Director of Beta NYC, and he's been working with

10 us and the CUNY Service Fellows to build new tools

11 making open data more useful to the unique needs of

12 Manhattans' 12 community boards, and I'm really happy

13 to report that the work has been successful.

14 Microsoft is working with us to even make the work

15 even more successful, and more transparent and more

16 accessible particular for the Board, and this

17 Saturday we are---we'll be announcing a 311 data

18 visualization tool that was built from the ground up

19 for community boards by—by the NYC Microsoft and the

20 students, and we certainly want to thank Mary

21 McCormick, Director of the Fund for City of New York

22 and the Alfred P. Sloan Foundation for making the

23 program and its success possible with their grants.

24 Thank you for giving me the opportunity to testify

25 today and I look forward because I know there are

2   changes that—these will be made to make these bills

3   even more appropriate for passage, and thank you, Mr.

4   Chair for all of your work. This is actually hard to

5   think that you will not be chairing this committee,

6   because you have done a phenomenal job. Thank you,

7   and I see my friend Madelyn—Council Member Palma over

8   there. It's nice to see you.

9   COUNCIL MEMBER PALMA: [off mic] Thank

10 you.

11   CHAIRPERSON VACCA: Thank you, thank you,

12 Gale and thank you again for being the pioneer and

13 the mover and shaker years ago for all that started

14 out that we've been able to build up.

15   GALE BREWER: All of us together.

16   CHAIRPERSON VACCA: All of us together.

17   GALE BREWER: Thank you very much.

18   CHAIRPERSON VACCA: Thank you, Gale

19 Brewer. Okay, se do have a panel: Mary Tobin,

20 Lexington Avenue, Brooklyn, Brownsville—representing

21 the Brownsville Partnership; John Kaehny, Reinvent

22 Albany, and Noel Hidalgo, Beta New York City. Please

23 come up. So, Mary, I'm going to call on you first.

24 Introduce yourself for the record and please give

25 your testimony.

2   MARY TOBIN: Good afternoon, Chair Vacca

3 and the members of the Technology Committee. My name

4 is Mary Tobin, and I have the honor of serving the

5 community of Brownsville, Brooklyn, as the Director

6 of the Brownsville Partnership, which is an

7 initiative of the Non-Profit Community Solutions.

8 Our community solutions we work very hard to end

9 homelessness and the conditions of poverty that

10 create it, and at the core of everything we do is

11 data. The Brownville Partnership's strategic plan

12 focuses on engaging residents and partner

13 organizations from many sectors immeasurably

14 improving the health, safety and economic prosperity

15 of Brownsville, which simply means we want to see

16 people find jobs, live in a safer environment and

17 attain a higher quality of life. How does data help

18 us do that? Well, we take existing neighborhood data

19 on everything from health to crime to education and

20 unemployment and we then analyze that data to

21 highlight any inefficiencies, gaps in -in existing

22 systems that serve the neighborhood we work in. We

23 use those insights t direct problem solving efforts

24 where they need to be applied the most. For example,

25 in my very own Brownsville in the Pathways

2 Initiative, we work on ways to increase the overall

3 number of people from Brownsville who gain

4 employment. Our early analysis of the Brownsville

5 data that we obtained from NYC Open Data, reveals

6 that there are three critical cohorts within the

7 general population of Brownsville who are facing the

8 most difficulty obtaining unemploy—obtaining

9 employment. Give the unique barriers to employment,

10 the specifically cohort studies: The Brownsville

11 Partnership Strategy has been to more actively engage

12 with in employers who work to employ these cohorts.

13 This very specific strategy has resulted in

14 meaningful conversations, more jobs and also very—

15 very meaningful strategizing sessions with public

16 agencies and other non-profits around how to improve

17 employment outcomes. This strategy also allows

18 organizations like mine with limited money, people

19 and time to direct our resources to the area to the

20 areas and the residents who need them the most. We

21 won't stop there as we are currently in partnership

22 with organizations using data to improve outcomes for

23 our most vulnerable population, children ages 0 to 3.

24 Because of the foresight of the leaders in this room,

25 and the work of the NYC Open Data Team, organizations

2 like mine are able to begin to pinpoint the most

3 critical issues in neighborhoods like Brownsville.

4 As we work together with the residents and community

5 based organizations to pool our own resources, test

6 our own strategies and solve our own problem. We

7 truly do believe that we can begin to move the needle

8 on poverty, but we aren't there yet, which is why

9 continuing to leverage programs like these is

10 critical to our work. For the work that my team does

11 everyday in a neighborhood that I've come to love as

12 my own, data is not just spreadsheets and charts. It

13 represents the life that deserves the opportunity to

14 thrive in conditions that other Americans enjoy every

15 single day. Thank you for your time and

16 consideration. Hope is truly inside of Brownsville.

17   CHAIRPERSON VACCA: I am very moved by

18 your testimony. I think it's fantastic. This is

19 exactly what we fight for. We fight to—we—we fight

20 for neighborhoods to use open data to empower the,

21 and you have used it, and you have been empowered,

22 and I want to thank you and your organization, but I

23 wish more people did that. I do think we still have

24 that education component--

25   MARY TOBIN: [interposing] Right.

2   CHAIRPERSON VACCA: --and an engagement

3 component--

4   MARY TOBIN: [interposing] Right.

5   CHAIRPERSON VACCA: --and an engagement

6 component that has to be really improved, but the

7 fact that you're there, and that you have that

8 enthusiasm, but that you're using the technology, and

9 we have fought so hard for at the Council to make

10 available to you--

11   MARY TOBIN: [interposing] Right.

12   CHAIRPERSON VACCA: --is—is rewarding to

13 me so.

14   MARY TOBIN: Absolutely. Thank you.

15 Thank you all.

16   CHAIRPERSON VACCA: You're great. Thank

17 you for being here.

18   MARY TOBIN: Thank you.

19   CHAIRPERSON VACCA: Thank you. Mr.

20 Caney.

21   JOHN KAEHNY: Thanks, thanks, Mary. I

22 haven't met you before, but that was great, and I

23 feel like I can leave now because this is—

24   MARY TOBIN: [interposing] [laughs]

25 Thank you.

2 JOHN KAEHNY: --really what we've been

3 fighting for a long time now.

4 MARY TOBIN: Absolutely. Thank you.

5 JOHN KAEHNY: Thank you. It's very

6 inspiring, and as—as thank go, you know, this is a

7 really big day for us to—to thank you, Mr. Chair

8 Jimmy Vacca. You know, when Mike Bloomberg left as

9 Mayor and Mayor de Blasio took over, no one was sure

10 where open data would go, but you've—you've kept it

11 going and pushed it ahead through, you know, highs

12 and lows, and we really, really appreciate your

13 tenacity and your intelligence and your forcefulness

14 and just making this a priority, and it's—it's

15 incredibly gratifying. So, thank you and from

16 everyone at the New York City Transparency Working

17 Group, all of the government groups that comprise our

18 coalition, you know, thanks for your effort, and for

19 keeping it going there, and it really matters and,

20 you know, I don't know who we're going to get as next

21 chair that can step into your shoes, but, you know,

22 for now thanks. The—I have some comments that are

23 both going to speak to the bills and to general

24 oversight, and your staff asked us to please not

25 forget oversight. So, we will talk a little bit

2 about oversight. Generally, just the state of open

3 data to us seems pretty good in New York City in part

4 because of you and you and people we haven't met who

5 are using the open data, and because we have a public

6 culture with civil society here within journalism,

7 within elected politics, community boards, community

8 based organizations of using open data in an

9 expectation that public data will be available in a

10 usable form by the public, and that expectation is

11 the most important legacy of the law and of your

12 persistent oversight. So, that's—that's good news,

13 and the other good news is that we have the funding

14 for the city's Open Data Team, who we think are great

15 an-and appreciate their work, and we don't take it

16 for granted that we'll always have that funding

17 available. So, I'm very pleased that you asked the

18 question about do you have the resources you need,

19 and I hope Council continues to fund the Open Data

20 Team going forward because without people keeping it

21 going, the technology will fall apart very, very

22 quickly. So, we're not forgetting that, and I also

23 wanted to touch on how important automation is, and

24 thank you for your questions about automating data

25 sets. We think in terms of sustaining open data as

2   well as the public culture and culture of oversight.

3   It's the automation of those data sets and the most

4   used data sets that is going to keep things going

5   when other things are not going so well

6   administratively in New York City. So, I want to

7   thank Albert who is not here, and his team and the

8   other folks at the Open Data Team for the automation

9   efforts, which we think is extremely, extremely

10 important. So, thank you for your—your asking

11 questions about that. I don't if folks here saw

12 yesterday 538.com, which is a big blog that does

13 statistical analysis, and they did a big story using

14 NYC City 311 data, and they shouted out to the Open

15 Data Portal, and they did a visual a visualization of

16 311 calls about Hurricane Sandy and the legacy of

17 Hurricane Sandy, but everyone should take a look at

18 it, and I thought it was an interesting example of

19 how Open Data and New York City's Open Data Law is

20 changing culture and perspectives, and just, you

21 know, leaking into all—all kinds of different areas.

22 Some specific oversight issues. Last year the—and

23 the year before that the Council passed seven Open

24 Data Law amendments, and they included new mandates

25 for standardizing addresses, and including data

2 dictionaries, and basically making the data more

3 usable.   We—we loved the amendments. We thought

4 they were great, and really helpful. The—but—but one

5 of the things that the amendments highlighted was the

6 fact that it is really, really hard to keep track of

7 what's happening with open data sets for the public,

8 when they're published, whether they're meeting the

9 open—the amendments or not, and just the information

10 about generally how well the agencies are complying

11 with the law. So, I'm glad to hear that Albert

12 Webber at DOITT is not having pushback from agencies,

13 but what we're seeing is fog when it comes to

14 understanding how well agencies are doing. So, of

15 course, we can see the data sets on the Open Data

16 Portal, but we're having a lot of trouble reconciling

17 the agency plans, their publishing plans with what

18 actually makes it all the way to the portal. So,

19 just a couple of stats that we pulled. They're—for

20 data dictionaries for instance, the city listed 1,648

21 data sets. 615 have—data sets have data

22 dictionaries, which is a lot more than we thought

23 would have. So, kudos to the Administration, and

24 1,033 don't. For address standardization, 350 data

25 sets were listed in the city's portal and 206 of

2 those are listed as having standardized addresses.

3 So, out of the roughly 1,700 published city data

4 sets, the city is confirming that 206 have

5 standardized addresses right now. One of the things

6 I would point to is that there are at least four

7 different spreadsheets scattered around the open data

8 portal for reporting on the status of different data

9 sets, and what we would like to see is one data set

10 that includes all of the city's public data--and I'll

11 get to this in more detail—all of the city's public

12 data and their status, whether it's published or

13 unpublished, when it's going to be scheduled, whether

14 it's been rescheduled, whether it's meeting the—the

15 address standardization mandate, whether it has a

16 data dictionary. All of that stuff in one data set

17 that the public can download and look at that's

18 machine-readable and easy to sort so that we can keep

19 track of what's going on because right now, the

20 information is spread over at least four different

21 sources, and it's incomplete. So, for instance, data

22 dictionaries, the city lists 1,648 data sets, the

23 status of their data dictionaries, but for address

24 and geospatial, they list 350 data sets. Why? We

25 don't know. So, rather than having separate scatted

2   data sets, we'd rather just see one big one where

3   you can look. Publishing. In the publishing status

4   of data sets, the core of the Open Data Law in 2012

5   was to define where the public data set was and then

6   to tell agencies that they had to start a methodical

7   way of publishing that data for the public. So, even

8   now, we still have trouble understanding what the

9   status is of planned data—of publishing dates, and

10   it's kind of a shifting sand for us. So, we drilled

11   down and we have a spreadsheet that I'll share with

12   you, and provide to the Administration as a

13   spreadsheet. We looked at six data sets, two street

14   tree data sets from Parks, the Fire Department's Fire

15   Incident, Construction Authority Funded Capacity

16   seats for schools, and HRA's Cash Assistance

17   Engagement data sets, and all those data sets were

18   planned for publication and scheduled in 2014, and of

19   those six data sets, four were never published, but

20   there was no update on what happened to those four.

21   Now, we picked those four at complete random. I mean

22   just, you know, throwing darts at the computer

23   screen, and the—- So, we hope that this is not a tip

24   of an iceberg, but does point to the fact that what

25   goes on and off of the—the data of the agency

2 publishing plans is—is very hard to track for the

3 public right now, and that's again why we just want

4 one website. I mean, pardon me, one data set that

5 just says when—when this thing is going to be

6 published and, you know, if it's rescheduled, when

7 and why just in one place. So—so that—so, you know,

8 good news overall, good culture, great stuff going

9 on, a great team, but some murky details here that

10 need to be sorted, and—and I'm going to get to some

11 specific recommendations in a moment. Overall, we're

12 really concerned about agency procrastination, and

13 the fact that there are so many data sets back loaded

14 to the second half of 2018 for publication. We count

15 102 data sets to be published in the second half of

16 2018, 70 of those in December of 2018. So, to us

17 that's like a kid, a high school kid promising to

18 turn in all their homework on the last day of school,

19 and we're having a little trouble, you know, buying

20 it right now. So, that segues into the—into the

21 comments on legislation. The 1528 open data FOIL

22 Amendment we strongly, strongly support that. It's a

23 small adjustment that helps, helps make that—that

24 bill more effective, and we partnered with existing

25 Local Law 7 of 2016 more effective. For those in the

2 audience who want to know what that secretive thing

3 is, it's—it just asks agencies to publish the list of

4 data sets they're using when they—by name when they

5 apply—when they reply to a FOIL request. On—on your

6 bigger bill, your joint with the Borough President,

7 we have some pretty extensive comments that revolve

8 around three areas, and we've already provided you

9 with some written comments, but basically we think

10 that—that this bill should have three goals:

11   1.   One to clarify that the Open Data

12 Law will continue past 2018, forever in perpetuity.

13 So we want to make sure that there's language in it

14 that makes it absolutely clear that the doc—that the—

15 that the Open Data Law continues, you know, past 2018

16 and past 2021. So that's one thing. Secondly, we

17 want to see strengthening of the mandate for the

18 agencies to continue to publish public data sets

19 forever. So, and I'll get tot that, and then lastly,

20 we want to see language that encourages and fosters

21 more data set automation. So, we think those are the

22 three big goals that should happen. Specific

23 recommendations. You know, this is complicated, but

24 we would like to see you keep the 2018 deadline for

25 publication. We are worried that the agencies are

2   going to feel that they're off the hook, and we're

3   not sure how to make them, you know, stick to their

4   schedule and publish the 2018 thing, and to that end,

5   you know, we—we think the 2021 extension though

6   there's a logic to it, might un—unfortunately feel—

7   give the agencies, you know, a homework extension

8   that they'll—they're never going to get to. So, we

9   think there's a way to—to get around that by keeping

10 the 2018 deadline in place, and then creating a new

11 requirement that—that new datasets be published

12 within 12 months or some time frame of when they're

13 identified. So, when a new data set is created,

14 whatever DOITT or the Administration feels is a

15 reasonable time, that data set should be, if it's

16 public and classified as public should then be under

17 a publication deadline. So, that's how we think

18 that—that this could work here, though we understand

19 what you're getting at. Just a couple more things

20 here. The big—big ask here is the—creating a new

21 mandate requiring a new status of all public data

22 sets. One big data set on the Open Data Portal, and

23 we thin that would help greatly, greatly increase

24 transparency and it would actually help the

25 Administration and the Council keep track of all the

2 data sets, ands all of the different mandates by

3 putting them in one place, one big long list and that

4 would be all public data sets whether they're

5 published or not published. Anything identified a

6 public data set, and then we could look at those and

7 we wouldn't have to go searching around, which is a

8 problem for us. We think that's especially important

9 because there's no public right of action or pardon

10 me, private right of action that allows the public to

11 sue agencies when they're not complying with the Open

12 Data Law, and so what you do and what we do is name

13 and shame the laggard agencies, and we're having

14 trouble understanding which agencies are laggard

15 right now because of—of the—the difficulty in

16 tracking the different data sets at least from the

17 public side. Therefor, we're hopeful that you add

18 the language to this bigger bill creating the single

19 what we're calling status of all public data sets,

20 which we think that would not be a burden at all on

21 the administration. In fact, would help them and is

22 something that we would think that they would want.

23 The—we have a list of the specific elements that we

24 think that all the data sets should have. Those are

25 starter, you know, scheduled publication, and current

2 scheduled publication, et cetera, et cetera, but I

3 think it's fairly apparent what we're asking for. We

4 would like to see another provision added to this

5 bill on automating data sets, and that would require

6 reporting fields and the status of all data sets,

7 data set let's say whether a data set is automated,

8 whether it can be automated, and then establishing a—

9 a schedule for automation of sorts. So, Albert

10 Webber said that 30 days or update--data sets that

11 update every month are reasonable targets for

12 automation. It sounds reasonable to us, but we would

13 like to see data automation codified and maybe

14 formalized a little bit. We think it's a little

15 loosey-goosey right now, and that it's—because it is

16 so important this needs to be track and in particular

17 the next chair who is not going to be as

18 knowledgeable as you about this is going to need to

19 have that kind of very clear reporting metrics. So

20 that would be another. And then, additionally, why a

21 data set cannot be automated. So, to summary—

22 summarize the—our take on the provisions in 1707, we

23 strongly support the review of the Technical

24 Standards Manual every two years. We support

25 changing the Compliance Plan deadline from July to

2 September. That's totally reasonable. We support

3 designated agency open data coordinators under the

4 law instead of voluntarily, and we support publishing

5 website portal analytics, but we would ask that you

6 consider not extending the publishing deadline to

7 2021 but keep the 2018 deadline in place with—with

8 additional language that clarifies that the Open Data

9 Law will continue in perpetuity, and then establish a

10 mandate for—for publishing new data sets the come

11 after that time. And then lastly, like Borough

12 President Brewer, we strongly support a private right

13 of action for the public to be able to sue agencies

14 who are not complying with the Open Data Law, and we

15 would note that there's only a handful of New York

16 City laws, a very tiny, tiny minority that don't have

17 a private right of action, and that the Open Data Law

18 is actually very unusual in that regard. So, thank

19 you for the extra time and much appreciated. Oh, I

20 should add that we support DOITT's proposal to add

21 licensing language about permissive licensing

22 language, we think is common sense to you. So, thank

23 you.

24   CHAIRPERSON VACCA: Thank you very much.

25 Thank you for all of your work. It's fantastic to

2 know that people like you will be continuing this

3 work, and your—your—your advocacy is very important

4 to us. Thank you. Noel.

5   NOEL HIDALGO: Hello. I want toe echo

6 everybody else's comments in regards to thanking you

7 personally as well as Borough President Brewer for

8 you steadfast leadership in open data and open

9 government and here in New York City. It's been

10 refreshing to go from one leadership to another

11 leadership and to be able to have that continuation.

12 So, thank you both for all of that work. I'd also

13 want to point that, you know, we are here because of

14 many elders. Beta NYC particular, a front gizmo to

15 the funds from the City of New York to NYPIRG to

16 Transportation Alternative who have been embracing

17 this data transparency movement for awhile and just

18 wanted to give a shout out and a thanks because the

19 elders of our community have been coming in and

20 providing us the type of feedback to help grow tools

21 like Board Stat that, you know, that Gale mentioned

22 and that the statement once again that Gale mentioned

23 around the world seeing New York City as an open data

24 leader. In all of my travels that I've had the—that

25 I've been fortunate to experience, time and time

2 again, it's how do—how did you et the Open Data Law

3 passed? How do you continue to have such a robust

4 program? How do you modernize your open data system

5 so that way you continue to stay at the forefront?

6 And, you know, across the United Nations—across the

7 United States, I—I get asked that question. Well,

8 this past week I had a chance to meet with somebody

9 from Valencia, Spain, and they were asking why, how—

10 how New York City be so far ahead, and here is a

11 great opportunity at this particular time to think

12 about how we can stay ahead. So, some of the

13 community things that—that DOITT mentioned

14 beforehand, the celebration of the fifth anniversary

15 of the Open Data Law. The—the inauguration--

16 inauguration of Open Data Week and the celebration of

17 international open data today converged this past

18 year, and we had over 350 people in attendance, and

19 this is the second year that we've hosted NYC School

20 of Data, and we think we've fine—have finally ironed

21 out a platform that enables government and community

22 to share their stories back and forth, and to keep—

23 empower each other. We've been able to in—in this

24 process include a number of curriculum items that

25 we're hoping to seed and experiment and then be able

2 to share through permissive licensing back to the

3 city so the city has a foundation as well as

4 community based organizations have a foundation to

5 further educate their different communities. In

6 regards to something that we've experimented with,

7 with the Parks Department, and City Planning this

8 year, and we're going to continue this weekend with

9 311 is we've started these things called data jams,

10 which enable one again the city government agencies

11 and—and partners, private partners to come in, and to

12 explore algorithms, to explore data set, kind of data

13 schemas, and this has been a great opportunity for us

14 to develop a model, and a replicable model that we're

15 happy to engage with the city to hopefully continue,

16 and we wouldn't be able to do NYC School of Data or

17 data jams if wasn't for the robust Open Data Program

18 that we have now, and to—and really those programs,

19 those programs, those civic engagement programs are a

20 testament to the mayoral leadership, and the agencies

21 adopting open data programs, and being open and

22 transparent about the—the fact that they want to

23 explore and build data driven decision making

24 programs. So, I wanted to commend those different

25 agencies that have been participating with us. In

2 regards to the Civic Innovation Fellows, a program

3 that we have partnered with the Manhattan Borough

4 President for the last few years, we've been able to

5 expose 25 undergraduate students to the city's

6 municipal infrastructure, and educate them on the

7 value of open data, and in the past two years we've

8 been able to move beyond open data, and we've been

9 able to map out community board district office

10 technical needs. This is something that we've

11 brought up a few times in some oversight hearing,

12 and—and that we're in the process of building Board

13 Stat, which is a tool that will simplify community

14 board access to 311 service request data, and we're

15 going to be launching it this weekend. What this

16 tool has exposed is continued need for vigilance

17 around data quality. There's some issues in regards

18 to 311, the 311 Geo Coder, and when you were asking a

19 question about where could resources be placed, we

20 really see that there's a hug opportunity. I didn't

21 writ this in here, but there's a huge opportunity to

22 improve Geo Support, which is a tool that so many

23 agencies use to translate B and BBLs, lat longs, the

24 addresses in general without having to pay thousands

25 and thousands of dollars to a private provider. And

2   what we have discovered is that if you are making a

3   mobile phone 311 service request, and you are in

4   Marble Hill, it will geo code your service request as

5   Community Board 8 Manhattan even though you are in

6   Manhattan, but you happen to be attached to the

7   Bronx, and your community Board is CB-8 in the Bronx.

8   And there is a few other data anomalies that we've

9   discovered, and we wish that Geo Support would have

10 the investment to bring it up to the 21st Century,

11 and hopefully be an open source stack that we can—the

12 community can help improve Geo Support. Let's see,

13 where to keep on going. Okay, to the legislation at

14 hand, so, okay. Hold on. I didn't—I'm trying to

15 summarize all of this very quickly. Five years into

16 the Open Data Program we still see quite a bity of

17 inequality. I can point to—I hate to use the

18 Department of Transportation as a punching bag, but

19 the—the two best examples that I have in regards to

20 the legislation that you introduced actually affect

21 the Department of Transportation, and so one of them

22 is the East River bicycle counts. So, every time a

23 cyclist goes over one of the East River bridges,

24 there's a little device that tracks and monitors the

25 bicycle going over it, and if I go to the Department

2 of Transportation's website, I can't get access to

3 that data. I only get a PDF, and when I actually

4 click onto where I can get access to the data on the

5 city's Open Data Portal, I get an Excel spreadsheet

6 that's been highly formatted on an annual basis and

7 doesn't conform to the Technical Standards Menu, and

8 it's frustrating because I know that this data set is

9 automated. I know that the DOT is getting that data

10 in an automated basis. There is very little concern

11 in regards to privacy violations with that particular

12 data, and so that data set should be something that

13 should be up on the data portal. It should be

14 automated, and—and it should be in access on a—on a

15 daily basis. Cities like Copenhagen and Amsterdam

16 they have gamified kind of bicycle cycling throughout

17 the city where there's little bollards that say how

18 many people have gone through this particular

19 intersection. And if you look at some reporting on

20 streets blog, it actually increase cycling activity,

21 and it's just one example of how we could turn the

22 city's open data into a more actionable environment,

23 and it's definitely something that we hope that the

24 modification to the Open Data Law around the

25 Technical Standards Manual would push the Department

2 of Transportation to make this data machine-readable

3 in a more real time way. The next example from the

4 DOT is around street closures, and I know that you as

5 the Chair of the Transportation Committee had to deal

6 with some of these issues, but as—as again as a

7 cyclist, I found myself this year going down Driggs

8 Avenue, and then all of a sudden a sudden coming

9 across the DOT maintenance team ripping apart the

10 bicycle lane that was there and there was no warning,

11 and for years, fore years and years and years, I've

12 used the example of street milling data as an

13 excellent example of if we could get this machine-

14 readable, we could be able to build an application

15 off of it, and just this spring I was complaining

16 about the fact that bicycle lanes seem to vanish out

17 from underneath me, and then I--all of a sudden I

18 have to wait six weeks if not two months to find out

19 when this bicycle lane would be replaced. And sure

20 enough, that happened again this year as I was

21 cycling down Driggs, and then as I was preparing my

22 written notes, I come to find out that DOT since the

23 spring has been putting that data on the Open Data

24 Portal. And to echo John's comments, you know, it's

25 very hard for us to track what data gets up on the

2 data portal when. We commend MODA and DOITT's

3 activities to get the data up on the data portal. We

4 wish that there was one conical (sic) where we could

5 see how data activity is—is being produced on the

6 city's Open Data Portal, and we would love for data

7 sets like this to be within the geospatial open data

8 standards, and we want this data set to comply with

9 that law, and sadly it's been automated, it's been—it

10 updates every single day, but it isn't complied—isn't

11 in line with Local Law 108 of 2015, the Geo—the

12 Geospatial Data Standards. So, I think that that

13 kind of addresses some of the other points, but one

14 concern that we have in regards to the legislation is

15 around the web portal analytics. The language right

16 now as it has been introduced says that—something

17 around people's location, and that the reports should

18 be around people's location, and we think that that

19 vague language could be strengthened to ensure and

20 protect people's privacy, which right you could

21 interpret it that everybody who visits the website

22 you can more or less report on their IP address and

23 you could identify like where exactly people are

24 coming from, and maybe we can work on some type of

25 terminology that put enough of the clarification to

2 protect people's privacy, particularly individual and

3 location privacy, and the last thing is we want to

4 echo the Private Right of Action as so many others

5 have—have stated. We really think that that would

6 cement the—an open data legacy for those of you who

7 have been working really, really hard at this. We

8 think that it would also provide the sustainable

9 framework for future activity, and lead—mark New York

10 City as open data leader—continue to make New York

11 City open data leader in the world. I think that's

12 it.

13   CHAIRPERSON VACCA: Wow. You gave us a

14 lot of things to chew on I guess, and all your

15 testimony was very technical and really helpful, and

16 we will take it back as we work on the bill, the

17 Borough President Brewer and myself. We want to come

18 up with the perfect product, of course. Whenever you

19 do legislation you want the perfect product, but we

20 want to have all the stakeholders like yourself at

21 the table. We thank you so much for coming. I want

22 to thank Council Member Palma who stayed for the

23 entire hearing today, which is fantastic. Thank you.

24 There being no further testimony, I thank this—I

25

2 thank-- Oh, do we have one more testimony? Oh, one

3 more person? Oh, do we have a slip.

4 SUMATA HARI HURASHURA: No.

5 CHAIRPERSON VACCA: Are you filling it

6 out now? Okay. We have one more person.

7 [background comments] [pause] Okay. Sumata.

8 SUMATA HARI HURASHURA: Hello, my name is

9 Sumata Hari Hurashura (sic) and I don't blame you for

10 not being able to pronounce that from my messy

11 handwriting.

12 CHAIRPERSON VACCA: Thank you.

13 SUMATA HARI HURASHURA: Sure.

14 CHAIRPERSON VACCA: You're from Astoria,

15 Queens?

16 SUMATA HARI HURASHURA: Yes, I live in

17 Astoria.

18 CHAIRPERSON VACCA: Okay.

19 SUMATA HARI HURASHURA: So, I am not here

20 representing anyone in particular. I very much

21 appreciate the Tweet that you sent out asking members

22 of the public to come and give their feedback on the

23 laws that you were considering today. I am an open

24 source software expert, and a programmer, and someone

25 who's—A New Yorker who is very much in favor of the

2 open data work that's been happening here over the

3 past year. So, I'm very grateful for both the bills

4 that are being discussed today to improve and expand

5 it, and make it more sustainable for city agencies.

6 So, I have three concerns, all of which have been

7 brought up by other people here. So, I'm just going

8 to second all of them just so you know that there's

9 at least one more person here who is also a

10 programmer who thinks all these are great ideas. One

11 is clarifying that there is still a 2018 deadline.

12 So, just as the Transparency Committee representative

13 mentioned so that we don't have the issue of these—

14 this bit of work being put off over and over again.

15 I very much agree with Mr. Webber about licensing. I

16 can say as someone who programs, looks for open data,

17 and tried to find openly licensed data so that I know

18 when I remix it, when I write presentations based on

19 it, when I add it to applications that I'm making, I

20 want to absolutely make sure that I'm in the clear

21 legally on the data that I'm using, and using

22 commonly agreed upon licenses that are attribution or

23 that are in keeping with the open definition are a

24 way that I can feel quickly that yes this sort of has

25 the brand name of openness stamped upon it. And so,

2   I would very much agree with Mr. Webber that having a

3   provision in here that says that license that are in

4   accord with the open definition would be fine. That—

5   that would be really great. Third, I want to agree

6   with Mr. Hidalgo and what he just said about the

7   website portal analytics section of this, and that's

8   actually specifically what I came to discuss. In the

9   section of 1707 that adds information, that adds a

10   provision about website portal analytics, the

11   language here is actually a little worrying from a

12   privacy standpoint. As Mr. Hidalgo mentioned, the

13   location from which a user accesses such portal is

14   vague enough that it could be interpreted to say all

15   the way down to their—their individual IP address,

16   which in many cases basically translates as a street

17   address, as a person's actual address where they

18   live. IP addresses in the Internet security

19   community are commonly understood now to be

20   practically PII, Private—Personally Identifying

21   Information to such an extent that they should be

22   treated with due respect and confidentiality, and

23   gotten—gotten rid of as soon as possible under

24   certain retention policies. So, clarifying that and

25   also clarifying that would be great. No—I think that

2 it might be a good idea to clarify this provision so

3 that in stead of page views, unique users, and the

4 location from which a user accesses such a portal,

5 specifying a little bit more about the parameters

6 that we're actually looking for would be. Location.

7 Do we care more about simply within or outside of New

8 York City or do we care about location down to the

9 borough level or the community board level or—or

10 something like because that's the kind of data that

11 potentially would be actionable by public data

12 advocates, agencies and so on as they manage and—and

13 incentivize particular kinds of work without posing a

14 privacy risk to people who are looking up open data

15 about sensitive topics. Page views great, number of

16 page views great. Unique users I don't want to be

17 identified uniquely in a public place for having

18 looked up a certain piece of—of data. So, number of

19 unique users perhaps is what this ought to say. So,

20 I wanted to mention these as concerns that I have and

21 a number of other privacy advocates as well would

22 have about this bill.

23   CHAIRPERSON VACCA: Very much appreciate

24 your testimony, and I share some of your concerns,

25 and we're going to take a look at it.

2   SUMATA HARI HURASHURA: Thank you very

3 much.

4   CHAIRPERSON VACCA: Thank you for coming.

5 I'm glad that the Tweet--[laughter] I'm glad somebody

6 reads my Tweets. [laughter] I Tweet all the time and

7 never--

8   SUMATA HARI HURASHURA: [interposing] So,

9 I'll—I'll—I'll speak a little bit more specifically.

10   CHAIRPERSON VACCA: Most times I see one

11 person or two people. I figure it's a relative of

12 mine somewhere I mean.

13   SUMATA HARI HURASHURA: [interposing] I'm

14 extremely--

15   CHAIRPERSON VACCA: I—I'm glad you res

16   SUMATA HARI HURASHURA: I'm extremely

17 excited about your Algorithmic—Algorithmic

18 Transparency bill, and so I had the--

19   CHAIRPERSON VACCA: If you're in favor, I

20 need you.

21   SUMATA HARI HURASHURA: I am and I-I

22 actually, you know about this, and so, I was keeping

23 an eye out to make sure that the—this was actually a

24 hearing that was available for public comment, and

25 just, you know, a soliciting to interesting people

2   saying interesting things. So, thank you very much

3   for clarifying in your Tweet that it was certainly

4   available for public comment.

5   CHAIRPERSON VACCA: Yeah, definitely, and

6   I'm glad you came, and the Algorithms Bill I think is

7   groundbreaking. I'm very much in favor of it, of

8   course. I introduced it, but--

9   SUMATA HARI HURASHURA: [interposing] I

10   imagine you did.

11   CHAIRPERSON VACCA: --next—next month I

12   hope you come back.

13   SUMATA HARI HURASHURA: I hope so, too.

14   Thank you.

15   CHAIRPERSON VACCA: I hope you come back

16   and thank you for your testimony.

17   SUMATA HARI HURASHURA: Thank you.

18   CHAIRPERSON VACCA: Thank you very much.

19   Okay, there being no further witnesses, I want to

20   thank everyone, Borough President, all the

21   stakeholders and the Council staff, of course, as

22   always. So, it's now 2:30 p.m. and the hearing is

23   now adjourned. [gavel]

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 ____ October 3, 2017 _____