CITY COUNCIL
CITY OF NEW YORK

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

TRANSCRIPT OF THE MINUTES

                    Of the

COMMITTEE ON TECHNOLOGY

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

                         January 24, 2017
                         Start: 10:01 a.m.
                         Recess: 10:56 a.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. Amen ra Mashariki, New York City's Chief
Analytics Officer, Open Platform Officer and
Director for Mayor's Office of Data and Analytics

Craig Campbell, Senior Advisor
Mayor's Office of Data Analytics

Dominic Mauro, Staff Attorney
Reinvent Albany
Member NYC Transparency Working Group

COMMITTEE ON TECHNOLOGY

2 [sound check, pause]

3 CHAIRPERSON VACCA: Everyone please take

4 your seats. How are you? Everyone please take your

5 seats. [pause] Good morning. My name is James Vacca

6 and I'm Chair of the Committee on Technology, and

7 we're here today to discuss three agencies'

8 compliance with the Open Data Law [coughs] and to

9 discuss the Mayor's Office of Data Analytics

10 examination and verification 2016 Findings Report.

11 Having open use with data can—that can be shared with

12 the public is a cornerstone of a transparent

13 government. Data can help New York City residents

14 improve their neighborhoods, advance research,

15 improve public safety and countless other benefits.

16 We passed the City's first Open Data Law in 2012

17 requiring eligible city data to be published on a

18 single web portal by 2018. We have continued to

19 improve law in subsequent years through a series of

20 amendments. One of those amendments is Local Law 8,

21 which is the subject of today's hearing. Local 8

22 requires MODA to create a plan to assess how select

23 agencies are complying with the Open Data Law and to

24 uncover data sets that might not have been disclosed

25 when agencies self-report. The agencies undergoing

2 MODA's examination this past year include the

3 Department of Sanitation, the Department of Housing,

4 Preservation and Development, and the Department of

5 Correction. The Department of Investigation affirms

6 that MODA's plan was either consistent with DOI's own

7 method of conducing investigations or that it

8 complied with recognized auditing standards. We

9 commend the Department of Information Technology and

10 MODA for helping New York City be a leader in making

11 public data available at the municipal level. Our

12 City government creates massive amounts of data each

13 day working with all of these agencies to ensure

14 their compliance with the Open Data Law, managing the

15 web of format agencies may have traditionally used to

16 publish data, and coordinating the publication of

17 this data on a single open data portal is no small

18 fete. Yet, DOITT and MODA have continued to make

19 strides. While we believe the agencies are in good

20 standing with the Open Data Law, there is room to

21 improve the process of examining the remaining six

22 agencies' compliance with the law in the next two

23 year as required by Local Law 8. We look forward to

24 getting additional insight for MODA on their first

25 year creating and implementing its examination and

2 verification plan. We also welcome advocates who

3 will testify today on ways to strengthen city agency

4 compliance with the Open Data Law. With the

5 continued joint efforts of the Council and all of

6 these stakeholders, we can ensure New York City

7 continues to serve as a model for open data. I've

8 been joined by my colleague Annabel Palma to my

9 right, and without further ado, I will now call the

10 Administration to give their testimony. I have to

11 swear you in. Please raise your right hand. Do you

12 affirm to tell the truth, the whole truth, and

13 nothing but the truth in your testimony before this

14 committee, and to respond honestly to Council Member

15 questions?

16   PANEL MEMBERS: [off mic] I do.

17   CHAIRPERSON VACCA: Okay, who would like

18 to lead off?

19   DR. AMEN RA MASHARIKI: [off mic] I will.

20   CHAIRPERSON VACCA: Okay, please

21 introduce yourself.

22   DR. AMEN RA MASHARIKI: [off mic] I will

23 turn my mic on. There we are.

24   CHAIRPERSON VACCA: Thank you.

25

2   DR. AMEN RA MASHARIKI: Good morning. My

3   name is Dr. Amen ra Mashariki. I'm New York City's

4   Chief Analytics Officer as well as the Open Platform

5   Officer and also the Director of the Mayor's Office

6   of Data and Analytics. I wanted to thank you very

7   much for inviting us here to testify on the work that

8   we did with the Examination and Verification Law.

9   Local Law 11 of 2012, also know as the Open Data Law,

10   requires city agencies to self-submit annual

11   compliance plans that lay out a schedule for

12   publishing all of their public data sets by 2019. In

13   the first three years of the Open Data program, this

14   method resulted in the publication of more than 1,500

15   data sets from 60 plus city agencies. Advocates, the

16   de Blasio Administration and City Council steered by

17   the leadership of the City Council—the Council

18   Technology Committee and Committee Chair Vacca felt

19   we could do even better. And in late 2015 and earl

20   2016 collaborated on several new laws designed to

21   provide additional mechanisms to ensure the

22   publication of every eligible data set. One of these

23   laws, Local Law 8 of 2016 was a new examination and

24   verification process. Its purpose is to improve

25   citywide compliance by creating a more systematic way

2 to locate datasets that may have inadvertently or

3 purposefully been excluded in agencies self-reported

4 open data plans. In January 2016, Mayor de Blasio

5 designated MODA as the office responsible for

6 carrying out the examinations and verifications. As

7 required by the law, MODA then created a plan for

8 conducting the process, which was approved by the

9 Department of Investigation Commissioner Mark Peters

10 in March 2016. I would like to express on behalf of

11 MODA and the Open Data Program our gratitude for the

12 opportunity today to speak to the quality of the 2016

13 Examination and Verification Report, our plans for

14 improving the process and our road map for better

15 citywide compliance with Open Data Law. Our

16 Examination and Verification plan consisted of three

17 parts. The first was a dataset questionnaire

18 intended to familiarize MODA with each agency's

19 current Open Data footprint, routine information

20 reporting requirements, data management systems, and

21 organizational structure. These questions served as

22 prompts for the Open Data coordinators at surveyed

23 agencies to verify that they have identified and

24 reported all eligible data sets in the agency's

25 highest priority information streams including FOIL

2 disclosures, data supporting performance indicators

3 and data included in reports required by legislation.

4 The second part was a certification letter modeled

5 after the all agency commissioners—the letter all

6 agency commissioners signed when they submitted their

7 first open data compliance plans in 2013. In this

8 statement, an agency commissioner or their designee

9 warrantied the accuracy and completeness of the

10 information provided in the data set questionnaire.

11 Finally, while users always have the option to

12 nominate data sets for publication on the Open Data

13 Portal, a specific window for public feedback

14 specifically invited for public to participate in the

15 examination and verification process. During the

16 firs two weeks of November, we invited feedback via

17 social media and alerts on the Open Data Portal. I

18 specified by the Local Law 8 the Department of

19 Correction, Department of Sanitation, and Department

20 Housing, Preservation and Development were included

21 in the 2016 examination and verification cycle. MODA

22 submitted its findings on these three agencies to the

23 Mayor, City Council, the surveyed agencies and the

24 public on December 1, 2016. In summary, we found that

25 all three agencies are in good standing with the Open

Data Law. Several data systems may have contained

eligible data sets, but warranted further review.

Three data set nominations were received during the

public feedback window, all of which were referred to

HPD. One was further referred to the Department of

Finance and another was determined to not be an

existing dataset. The third was a request for daily

updates to HPD's emergency repair charges dataset,

which was reported as under further review, but since

been added to HPD's Open Data Compliance Plan. All

three agencies cooperated fully and provided valuable

feedback during our inaugural round of the

examination and verification. We would in particular

like to thank the three agencies' Open Data

Coordinators along with Commissioners Bean, Garcia

and Ponte for being forthcoming with our requests and

ensuring that this process ran smoothly and on time.

The law also requires MODA to report any deviations

to the plan we submitted to DOI. While we followed

the DOI approved plan precisely, we did encounter

challenges carrying out of the—carrying out the

verification workload that we had intended. We

explained these difficulties in detail in the report.

These challenges include issues determining the

2 eligibility of raw data that may have never before

3 been extracted as a dataset from its transactional

4 system, and conceptual questions on the scope of

5 public datasets that were not immediately resolved by

6 the definition provided in Local Law 11 of 2012. In

7 the report, we proposed ways to meet these challenges

8 in our recommendations on better citywide compliance.

9 In particular, we recommend that agencies should look

10 for ways to ensure that newly collected data can

11 easily be published as open data. For example, the

12 contract procuring Citi Bikes' technical

13 infrastructure included a provision mandating that

14 the data collected be made public. Because this open

15 data requirement was built into the technology itself

16 the ride data published by Citi Bike is easier to

17 access, easier to use, and in high demand. Two, the

18 open data team should take more steps to empower Open

19 Data Coordinators. Earlier this month, we surveyed

20 all of the Open Data Coordinators to better

21 understand their roles within their agencies, the

22 challenges they face and opportunity for improved

23 coordination. The results are informing our ODC

24 engagement and recruitment strategy, which begins

25 with convening all of the Open Data Coordinators in

2   City Hall in mid February. And lastly, MODA should

3   improve the Examination and Verification Plan. One

4   of our strategies is around continuous improvement.

5   For the 2017 strategy, the examination and

6   verification process will coincide with the Open Data

7   reporting cycle that culminates in the publication of

8   the Annual Open Data Plan on July 15th as mandated by

9   Local Law 11 of 12—of 2012. Aligning these

10 processes, will make agencies' overall reporting

11 obligations less resource intensive and the MODA

12 process more efficient. In the meantime, we are

13 conducting paired research and stakeholder interviews

14 to produce guidelines that will make it easier for

15 agencies to complete the examination and verification

16 process, and self-verify their compliance. We intend

17 to make these guidelines public when we publish the

18 2017 Examination and Verification Report. In many

19 municipalities, open data is an IT policy. In

20 others, it's an executive order. The business model

21 for open data elsewhere is based on the direct

22 returns it provides to city government like reducing

23 the cost of FOIL data disclosures, serving as a

24 resource to internal and industry app developers, and

25 representing a gesture of government transparency.

2   All holds true for New York City, but New York is

3   unique in that—in that here open data is the law.

4   Our vision for open data for all is the idea that

5   open data belongs to New Yorkers naturally followed

6   from the fact that elected officials via the

7   constituents they represent were so committed to this

8   idea that they added it to the New York City

9   Administrative Code. We believe that the fundamental

10 promise of open data is that it avails public

11 information that gives people the knowledge to take

12 action on their own behalf. As a policy, it is

13 similar to the Freedom of Information Law, which

14 guarantees that documents and information maintained

15 by government be disclosed when requested by the

16 public. But while FOIL is intended to deliver

17 precise answers to specific questions or disclosed

18 documents in a fixed form, data is more liquid. It

19 is for this reason that in addition to the report we

20 included itemized results of the agency dataset

21 inventory surveys in the—in the Examination and

22 Verification 2016 Results Excel Workbook that came

23 along with the survey. This file is meant to provide

24 users with a snapshot of the technical ecosystem of

25 the agency, and a better understanding of how data

2 gets from an agency data source to a usable data set

3 in the Open Data Portal. We encourage users to take

4 this information to request new data sets for

5 publication with the data set nomination feature,

6 which guarantees a formal review and timely response

7 by Local Law 109 of 2015. These requests are helpful

8 in supplying parameters that guide decisions on how

9 to represent data that may previously not have had

10 any bound and fixed form. They also signal public

11 interests in an agency's data, which according to the

12 prioritization parameters name in Local Law 11 of

13 2012, instantly escalates the urgency of making a

14 determination on how to publish an appropriate data

15 set. A demonstration of public interest is crucial

16 to the ongoing success of the Open Data Initiative

17 and it's compliance with all t he associated laws.

18 As it makes agencies more aware of the data they have

19 and maintain. This in turn spurs better upkeep and

20 disclosure to the public and the cycle continues. It

21 is also incumbent on the Open Data Team to cultivate

22 public demand. For this reason, when we joined the

23 New York City Open Data Community to celebrate the

24 five year anniversary of the Open Data Law this

25 month, we are launching a new website that makes it

2 easier than ever to request new datasets for

3 inclusion on the Open Data Portal. In addition to a

4 better nominations interface, the new cycle will

5 include a feature that allows users to track the

6 status of their requests. The new website and the

7 engagement we will conduct as part of rolling it out,

8 will also make it easier to access resources to learn

9 more about how to use open data, track our progress

10 on implementing all of the relevant Open Data laws

11 and get in touch with the Open Data Team on anything

12 from technical inquiries to partnership

13 opportunities. As we improve the examination and

14 verification over the next year, we will also

15 leverage the process to ensure compliance with all of

16 the Open Data Laws including the provision of data

17 dictionaries and common geospatial fields in all data

18 sets as well as advancing our strategic goals to

19 improve data quality and proactively engage the

20 public on new data set releases. The legislative

21 mandates increasing user base and growth of open data

22 culture in city agencies, put us in a strong place

23 moving into the law's final years of industry

24 compliance. We look forward to leveraging the

25 lessons we've learned and working with the Council to

2 plan for what comes after December 31, 2018. Thank

3 you very much.

4   CHAIRPERSON VACCA: Thank you very much

5 for your testimony. We've been joined by Council

6 Member Barry Grodenchik, my colleague to my right. I

7 did have some questions, and let me first start with

8 you talk about leveraging the lessons we've learned

9 at the end, and you're talking about legislative

10 mandates. My question is do you think that your

11 agency will be recommending revisions to the Local

12 Law? Are there additions to it or modifications that

13 you think you may need based on your experience?

14 Because then I'd—I'd like to know how we in the

15 Council can be helpful.

16   DR. AMEN RA MASHARIKI: So the way we're

17 looking at it is that we've built the process such

18 that it can build on top of itself, and so I would be

19 in a better position to answer that question after a

20 second round because we've identified smarter ways

21 and more efficient ways to go about doing those

22 spaces. As I mentioned, instead of doing this after

23 we release the Open Data Report for 2017, we're going

24 to do the examination and verification process during

25 the work that we do to ultimately release the Open

2 Data Report in July. And so, we believe that that's

3 going to give us a great opportunity to engage the

4 ODC's to do an even more robust and efficient

5 examination and verification. So I think after this

6 round, we'll begin to see where we may or may not

7 need additional help.

8   CHAIRPERSON VACCA: Well, you're talking

9 about working with several agencies. You mentioned I

10 think four: HPD, Sanitation and what is the one--?

11   DR. AMEN RA MASHARIKI: DOC.

12   CHAIRPERSON VACCA: Excuse me?

13   DR. AMEN RA MASHARIKI: DOC.

14   CHAIRPERSON VACCA: DOC and Sanitation

15 and HPD. When are you bringing more on line? Do you

16 have a timetable that you can share?

17   DR. AMEN RA MASHARIKI: We're staring

18 immediately in February to begin the—the—the—the

19 following three.

20   CHAIRPERSON VACCA: What are the next

21 three?

22   CRAIG CAMPBELL: [off mic]

23   CHAIRPERSON VACCA: Introduce yourself

24 and yes.

25

2 CRAIG CAMPBELL: [on mic] My name is

3 Craig Campbell. I'm the Senior Advisor for the

4 Mayor's Office of Data Analytics. I work closely on

5 the Examination and Verification process. The next

6 three agencies, which we will pick up the process

7 this May are FDNY, the Department of Buildings and

8 DEP.

9 CHAIRPERSON VACCA: You're aware that

10 this committee had raised the question as to when the

11 Police Department would online, and my concern is

12 that we have pushed that too far into the future.

13 Tell me where we stand with that, and you can in

14 cooperation there, and if it's something that we can

15 look forward to in the next several months.

16 DR. AMEN RA MASHARIKI: So in general we

17 haven't been getting great cooperation from NYPD with

18 regards to open data. We're following the strategy

19 laid out in the law with respects to the agencies to

20 engage, but we can begin having a discussion about

21 adding them to the list.

22 CRAIG CAMPBELL: The NYPD was not one of

23 the specific agencies named in the law, but as part

24 of this continued strategy, we do plan to leverage

25

2 the lessons learned from the minimum amount of

3 agencies to everyone that we work with.

4   CHAIRPERSON VACCA: So you're saying that

5 the Police Department is not ready at this point.

6 You're not ready at this point?

7   DR. AMEN RA MASHARIKI: No, that's not

8 what we're saying. We have a process in place to do

9 examin—Examination of Verification. Based of the

10 resources that we have, we've identified a schedule

11 that matches the requirements laid out in the law.

12 And so that's what we've—we've put together. We can

13 being looking into—as Craig just mentioned, we can

14 begin looking into ramping up other agencies. But

15 again, like I said, the way we're trying to do this

16 is not to bifurcate the—the two processes, which is

17 the Open Data Plan process and Examination of

18 Verification. We're actually trying to bring those

19 together such that the process that we go through to

20 engage the ODCs and the agencies with which to

21 release the yearly Open Data Plan, is built into the

22 Examination of Verification process just that it's a

23 smarter process to engage all agencies, and the best

24 way to identify data sets that exist within the

25 agency.

2   CHAIRPERSON VACCA: Will you discuss the

3 new committee, the need to have an Open Data

4 Coordinator designated within city agencies? Whether

5 it was a person—person specifically hired for that

6 purpose or whether it was someone else on staff who

7 served in that capacity. Tell me where we are with

8 that? Do we have agencies now who have designated

9 Open Data Coordinators and do we have some others

10 that have not? Is there a policy that's going to be

11 put forth from the Administration?

12   DR. AMEN RA MASHARIKI: Absolutely. We

13 have over 80% agencies that have Open Data

14 Coordinators. What you'll see is turnover in terms

15 of people leaving their positions, leaving the agency

16 and so on and so forth. So that accounts for a lot

17 of times we're missing—some—some agencies won't have

18 a designated ODC, but we're working on identifying

19 all of those ODCs. Like we mentioned, we're going to

20 have a kick-off meeting in February at City Hall to

21 bring all of the ODCs together and we're going to

22 pull that together.

23   CHAIRPERSON VACCA: I'm interested in

24 making sure that the Open Data Coordinators are

25 somehow listed on the city website. Will those

2   agencies, Open Data Coordinators be listed on your

3   website or will they be under the individual agency

4   websites? I want people to know who the Open Data

5   Coordinator isn't. How do we best do that?

6   DR. AMEN RA MASHARIKI: I think we haven't

7   made that determination yet. One of the things that

8   we've seen with Open Data Coordinators is that they

9   could be virtually anyone in the agency. It could be

10   the Deputy General Counsel. It could be an IT

11   person. It could be a newly minted staffer. And so,

12   often times this is—has nothing—it may have nothing

13   to do with their day-to-day job. They've been

14   assigned this, and so what we want to do is begin to—

15   as we mentioned, we're engaging ODCs in a very sort

16   of aggressive way. But we surveyed the ODCs. We did

17   a really thorough survey with the ODCs, and we're

18   trying to understand exactly the best way to allow

19   for the ODCs to be responsive to the agency, but also

20   not to lose them. I'd put them in a position where

21   they'd sort of inundate with external requests. So

22   we're—we're happy to discuss any suggestion that you

23   may have, but we're thinking—we're thinking through

24   the best strategy.

25

2   CRAIG CAMPBELL: And part of the—we're

3   trying to bring the legislative mandate of Local Law

4   109, which requires timely responses to data set

5   requests with the new technology that we'll be

6   rolling out as pat of the website release in early

7   March to ensure that—that when someone submits a

8   comment or a complaint that it is fitting to the ODC

9   and that they can see how their—the status of that

10 comment, and ensure that they have a complete

11 response.

12   CHAIRPERSON VACCA: Today, we're talking

13 about DOC, Sanitation and HPD. Do they have

14 designated Open Data Coordinators?

15   DR. AMEN RA MASHARIKI: Yes.

16   CHAIRPERSON VACCA: They do?

17   DR. AMEN RA MASHARIKI: They do.

18   CHAIRPERSON VACCA: Okay. For the sake

19 of transparency, I'm interested in having somewhere a

20 record of who those Data Coordinators are at least

21 for those three agencies, but then for the next four

22 I think you indicated in February Fire and DEP, and

23 those agencies. I think if we can't do the entire

24 city at this point because some agencies are in

25 process, I think that we have an obligation for

2   transparency with the agencies that are mandated by

3   law as we need more. At least that has to be posted

4   somewhere. I would think your website is the most

5   logical place for the—within the agency website

6   somewhere that people—where people can see it, and

7   I'd like that to be prioritized.

8   DR. AMEN RA MASHARIKI: Absolutely.

9   CHAIRPERSON VACCA: Now, we spoke at this

10   committee about non-mayoral agencies and their

11   compliance, and I want an update on that. I'm

12   thinking of the Housing Authority, NYCHA, Health and

13   Hospitals Corporation. Where are we with those

14   agencies complying, and do you have relationships and

15   coordinators established there and what—how are they

16   progressing?

17   DR. AMEN RA MASHARIKI: We do have

18   relationships at those organizations. At NYCHA, you

19   know, last year's report has released data onto the

20   portal, and is working with us in a—a productive way.

21   H&H we have a relationship as well. H&H does have

22   some data on the Open Data Portal, and we're

23   continuing to engage them as well. I think there are

24   some conversations that need to be had in terms of

25   the requirement for them to release, but we've

2 fostered a strong relationship with their data

3 people, and they've been extremely forthcoming and

4 helpful.

5   CHAIRPERSON VACCA: Do you have an idea

6 of how many people in our city use open data? How do

7 I know how many people use it, and that's a multi—

8 multi-part question. How many people use it citywide

9 among our eight million plus residents, and—and the

10 other question would be: Are agencies actually using

11 it to conduct their business everyday? Do they know

12 how to use it? Are they using it? These are

13 important questions, and how do we get a gauge on how

14 many users we have?

15   DR. AMEN RA MASHARIKI: So those are—

16 those are all very important questions. Let me

17 answer the last one first in terms of our agencies

18 using it. The answer is yet. We've worked with

19 certain agencies to put their agency data online

20 specifically because their technology leadership

21 reached out and said, Hey, we've got other agencies

22 pinging us about this data set, which we normally

23 would put on the hard drive or put on a disk and get

24 to them, but we would like to put this onto the Open

25 Data Portals, and then appoint that agency to use it.

2   So there's not only intra but interagency usage of

3   data. I know specifically we've spoken with DOE, and

4   DOE puts the data sets from specific departments

5   online so that other departments within DOE have

6   access to it, and can use it. DOE just did a

7   hackathon where they released open data to do an

8   analysis on enrollment—student enrollment and so on

9   and so forth. So we've seen—our Parks Department has

10 done a great job of really getting out in front and

11 showing how they can release the data sets that not

12 only other agencies can use it, but people within

13 Parks Department use it as—use their own data as

14 well. So, we've seen really strong examples of

15 agencies utilizing open data. With respects to the

16 question of who's using open data. So we've actually

17 journeyed along that question starting in the summer.

18 We started with trying to understand the breadth of

19 all of the comments that have come in. So what we

20 did was we did a full breadth from the day the Open

21 Data Portal was stood up, to the current day, which

22 was the middle of summer, we looked at every single

23 comment, and we did a full typology to understand

24 exactly who was commenting and what questions were

25 they asking about, and so we did a full assessment

2   that we're happy to share on who actually commented.

3   So it's not just who went to the portal, but who went

4   to the portal and was looking to engage in some way,

5   shape or form who had questions and so on and so

6   forth. So we did a full assessment of that. What we

7   did find is that to count to truly sort of graph as

8   MODA, we're very interested in data and the truth

9   within the data. And so to truly count and do an

10   assessment of how many users are—are with access to

11   portal, it's—we have to do sort of a—a full analysis

12   of different data sets to get theirs. It's not a—a

13   trivial—a non-trivial sort of process. It's very sort

14   of complex to kind of-- So, product counts one way,

15   agencies count another way, and so we have to pull

16   all that data together, and so we've started that

17   process.

18   CRAIG CAMPBELL: And I'll follow up in

19   terms of volume that we report the statistics each

20   year in July 15th Report. Last year the most visited

21   dataset in the portal had five million views, and

22   with all of the mobile apps and other websites, that

23   full data directly from the Open Data Portal there

24   are billions of those transactions underway.

25

2   CHAIRPERSON VACCA: Okay. Let me—but I

3 have more questions, too, but we have Council Member

4 Borelli who just joined us, and we welcome him. So

5 we don't have actual numbers right now. We don't

6 have actual number of users?

7   DR. AMEN RA MASHARIKI: No, not the

8 actual number of uses right now.

9   CHAIRPERSON VACCA: But when will you be

10 getting actual numbers? Because then once we get

11 actual numbers we can decide whether or not we have

12 to do more outreach and more engagement of community

13 groups. Whether or not agencies really are making

14 use of it as they should to document their budget

15 requests, to document the level of complaints that

16 they receive and—and actions they take. So I'm

17 interested in numbers because we had a slow start

18 here, you know. I'm—I'm satisfied now that we're

19 progressing, but we had a slow start in the

20 beginning. So when we have this kind of information?

21   DR. AMEN RA MASHARIKI: So we're working

22 on those, and we'll get those to you in short order.

23 What I will say is without seeing the actual numbers,

24 one thing we're clear on is that engagement not only

25 needs to happen but needs to happen aggressively. No

2 matter what the number says, it's not nearly reaching

3 the number that we want to get to, which is all New

4 Yorkers so--

5   CHAIRPERSON VACCA: [interposing] But—

6 but—but then we need the strategy for New Yorkers.

7   DR. AMEN RA MASHARIKI: That's right.

8 That's right.

9   CHAIRPERSON VACCA: How do we engage, and

10 what do we do to publicize the use of the Open Data

11 Portal, and all the information it contains? So I

12 look forward to seeing that strategy soon.

13   DR. AMEN RA MASHARIKI: Oh, you will.

14   CHAIRPERSON VACCA: Okay. Now, you said

15 here in your testimony new website, and you were

16 talking about the new site would include a feature

17 that allows users to track the status of their

18 requests. So explain that to. How would they track

19 the status of their request?

20   DR. AMEN RA MASHARIKI: So we're putting

21 in—we've built a workflow that allows for anyone who

22 comes to the website to put in a request any comment,

23 and we're putting in a workflow technology on the

24 back end that manages that. Not only points that to

25 the appropriate agency to respond through MODA and

2 DOITT, but then also we're tracking sort of where

3 that's going, and where we're going to share that

4 back to the users.

5   CHAIRPERSON VACCA: Are we talking about

6 requests for data, or are we talking about a request

7 that they make citizen complaint? If they make a

8 complaint to 311 let's say. This does not involve

9 that. This only involves the data.

10   DR. AMEN RA MASHARIKI: Correct.

11   CHAIRPERSON VACCA: Okay, you mentioned

12 Freedom of Information requests. Has there—has there

13 been a reduction in Freedom of Information requests

14 based on the Open Data legislation?

15   DR. AMEN RA MASHARIKI: I don't—I don't

16 think we're—we've tracked that metric, but we can

17 look into that.

18   CHAIRPERSON VACCA: Yes, you did mention

19 that in your testimony, and I'd be interested in

20 knowing the impact it's had on FOIL requests as well.

21   CRAIG CAMPBELL: We have anecdotal

22 evidence of that from the agencies that we work with

23 one of which is TLC. Recently it move to put all of

24 its for-hire vehicles drivers' dataset on the Open

25 Data Portal, which reduced the one of just like

2   portal pleasures where people would have to come

3   within, you know, these dates, but they'd load all

4   the data onto parts of data closing on open data.

5   (sic) And that has reduced a lot of common efforts

6   for them, and our managing those requests.

7   CHAIRPERSON VACCA: Okay, any questions

8   from the Council Members? Yes, Annabel Palma.

9   COUNCIL MEMBER PALMA: Thank you, Mr.

10   Chair. Thank you for raising most—most of the

11   questions and concerns that the members of this

12   committee have around this issue. I'm interested in

13   knowing have any of the agencies reported any

14   security challenges when dealing with publishing the

15   data or storing the data or admit perhaps why they

16   might not be up to par when complying with the Local

17   Law?

18   DR. AMEN RA MASHARIKI: None in—in terms

19   of any mishaps or any challenges, but the process

20   that MODA and DOITT engages each agency on does walk

21   each agency through sort of a security and privacy

22   strategy and process. So we engage the agencies

23   heavily, and there's been no sort of as you've

24   pointed mishaps or challenges there.

25

2   COUNCIL MEMBER PALMA: In—in terms of the

3 Department of Education and the Department of

4 Homeless Services, we know where they are in terms of

5 compliance, and when MODA will have them come on—come

6 on line for this—for the publishing or the starting

7 of their—their—of them. (sic)

8   DR. AMEN RA MASHARIKI: So I don't have—

9 we've published the Open Data Report and the Open

10 Data Plan. We can get to you the specific details on

11 DOE and DHS' data sets, but as I know, they're—

12 they're compliant.

13   COUNCIL MEMBER PALMA: Thank you. Thank

14 you, Mr. Chair.

15   COUNCIL MEMBER GRODENCHIK: [off mic] Mr.

16 Chair.

17   CHAIRPERSON VACCA: Yes, Council Member

18 Grodenchik.

19   COUNCIL MEMBER GRODENCHIK: Thank you

20 very much. In your testimony this morning, Doctor,

21 you had mentioned the high demand about Citi Bike.

22 Can you tell me what high demand is in your opinion

23 on the—on the city web? Are we talking hundreds of

24 thousands or it's millions or tens of thousands?

25

2   DR. AMEN RA MASHARIKI: I don't have a

3   specific number.

4   COUNCIL MEMBER GRODENCHIK: Do you keep—

5   who keeps track of how many hits the city agencies

6   get on their websites?

7   DR. AMEN RA MASHARIKI: My assumption

8   would be the city agency.

9   COUNCIL MEMBER GRODENCHIK: It would be

10 the city agency who's responsible. Do you have any

11 idea how many people access the city website every

12 day or every week or every month or--?

13   DR. AMEN RA MASHARIKI: No.

14   COUNCIL MEMBER GRODENCHIK: Who's—is

15 there somebody responsible for keeping that data?

16   DR. AMEN RA MASHARIKI: I suspect so. Not

17 in--

18   COUNCIL MEMBER GRODENCHIK: [interposing]

19 Not in your office.

20   DR. AMEN RA MASHARIKI: --in my office

21 then.

22   COUNCIL MEMBER GRODENCHIK: Okay.

23 Alright, thank you, Mr. Chairman.

24   CHAIRPERSON VACCA: Now, one of the

25 recommendations of MODA's 2016 Findings Report was to

2   create clear guidelines on the def—and definitions of

3   data and data set, and then to determine what data

4   set is public and private. Now, these—these

5   definitions were not established prior to the

6   December deadline. Why—why was that? They're—

7   they're fundamental to the law to be complied with.

8   DR. AMEN RA MASHARIKI: Correct. So the—

9   the definitions. We went into the Examination and

10   Verification process with an established definition

11   based on when the Local Law 11-2012 was passed. What

12   we saw, which shows the—the—the strength of the

13   Examination and Verification Law, what we saw was

14   that we actually waited to add a few more or—or—or at

15   least investigate and maybe expand the list that we

16   had that identified what is or what isn't a public

17   data set. And so, it's not that we didn't have

18   anything going in, it's that's this process has

19   helped us think a little bit smarter about what

20   should or shouldn't be a public data set. More

21   questions came up because of the process--the

22   Examination and Verification process, which the way

23   we saw it was that—that the process is working if

24   it's costing us. We began to think about what's

25   public, what's private, and actually to be very

2   honest, as we begun to really dig into this concept

3   of privacy and private data, I think that's a—a

4   conversation that should consistently be had, and we

5   should consistently look into that.

6   CHAIRPERSON VACCA: I know I avoided

7   using this term in the legislation because I—I

8   produced this legislation, and I didn't want to use

9   the term audit. We had a discussion with the

10   Administration about using the word audit, but an

11   audit is what this basically is, an examination and

12   investigation of an agency by your agency to make

13   sure that there is compliance with the Local Law.

14   So, my question to you is when there's an audit there

15   is a report agency by agency. Is there a report

16   generated on compliance agency by agency?

17   DR. AMEN RA MASHARIKI: That—that we

18   release?

19   CHAIRPERSON VACCA: No.

20   DR. AMEN RA MASHARIKI: Yes.

21   CHAIRPERSON VACCA: That you release?

22   DR. AMEN RA MASHARIKI: Correct.

23   CHAIRPERSON VACCA: There is, and that

24   report is similar to an audit in that it basically

25   says what you found, and makes recommendations on

2   what to do next, and how to improve and what was not

3   complied with that should be complied with?

4   DR. AMEN RA MASHARIKI: Correct.

5   CHAIRPERSON VACCA: And the agency does

6   as they do in an audit, the agency sends you a letter

7   agreeing to the audit that they will comply?

8   DR. AMEN RA MASHARIKI: Correct.

9   CHAIRPERSON VACCA: Okay. I just wanted

10   to make sure. So it is an audit, but we didn't term

11   it that far, but I—I—I understand. Okay, any further

12   questions from the committee. Council Member, any—

13   any further questions? Hearing none, I will thank

14   you for your testimony, and we have a panel so I'll

15   call them next. Thank you all.

16   CHAIRPERSON VACCA: Dominic Maurro

17   That's the panel, one person. [background comments]

18   That's okay. Reinvent Albany. [background comments,

19   pause] Okay, please introduce yourself and proceed.

20   DOMINIC MAURO: Good morning Chairman

21   Vacca and members of the Technology Committee. I'm

22   Dominic Mauro, Staff Attorney of Reinvent Albany and

23   a member of New York City Transparency Working Group.

24   I want to start by thanking you Chairman Vacca and

25   the members of this committee and the City Council

2   for your continued commitment to oversight hearings

3   for the Open Data Law. Their ongoing an energetic

4   support for open data has made New York City a global

5   leader in open data, and it's hugely encouraging to

6   open data advocates inside and outside of government.

7   Also, our sincere thanks to Minnie Tarlow, the

8   Director of the Mayor's Office of Operations, and

9   DOITT Commissioner Anne Roest who have helped staff

10   up the City's Open Data Team, and have dedicated more

11   time to open data issues. We also thank the Open

12   Data Audit Team at MODA and DOITT for their earnest

13   and professional work carrying out this first ever

14   Open Data Audit. We are extremely pleased to see the

15   Administration comply with Local Law 8 of 2016 in a

16   timely and serious way. We have three brief comments

17   on the agency Open Data Audit.

18   First, the Administration's Open Data

19   Team exceeded our expectations and gathered and

20   shared with the public a great deal of useful

21   insights.

22   Second, and we find this odd given the

23   overall high quality of the audit. The Open Data

24   Team declared all three agencies in compliance with

25   the Open Data Law, but the evidence they gathered

2   raises questions about whether the agencies are

3   complying with the Open Data Law. Specifically, the

4   audits found 20 public data sets, which the three

5   agencies have not scheduled for publication on the

6   Open Data Portal. The Open Data Law requires all

7   public data sets to be published by the end of 2018.

8   Department of Sanitation should be considered out of

9   compliance with the Open Data Law until it puts the

10 14 public data sets on a schedule for publishing

11 before the end of 2018. And the same goes for six

12 public data sets identified a Corrections and HPD.

13   Third, and more positively, the Open Data

14 Team lists a series of forward looking

15 recommendations on page 5. We strongly endorse all

16 eight of these specific recommendations, and hope

17 that City Council and public stakeholders are invited

18 to engage in the process of implementing them. So

19 the Open Data Team's audit was thoughtful and

20 included a number of useful features. First, it

21 describes the data sets, which are used to calculate

22 each MMR indicator for over 100 indicators. It

23 inventories each agency's technical systems with more

24 than 20 users, organized by agency program, and it

25 lists the agency personnel consulted for the

2 expertise and their respective agency's data assets.

3 And finally, it examined agency FOIL logs for

4 repeated requests for public data sets although the

5 agencies did not identify any data sets to be

6 published. However, we do have serious concerns

7 about two parts of the report.

8   First, as I mentioned before, we do not

9 understand how the agencies can be considered in

10 compliance with the Open Data Law when they have no

11 plan to publish the 20 public data sets identified by

12 the Open Data Team. According to the Open Data Law,

13 a data set is either public or not public. The Open

14 Data Team and the agency have to decide, and they

15 have to explain why a public data set is not part of

16 the agency's compliance frame. The Open Data Team

17 explains that these 20 public data sets are clearly

18 public or clearly private. I'm sorry, they are not

19 clearly public or clearly private. They are less

20 definitive and they require further investigation.

21 That's the footnote on page 4, but the main purpose

22 of the Local Law 8 Audit is to tell the world how

23 many public data sets an agency has published, how

24 many they've scheduled for publishing and how many

25 public data sets have not, but should be scheduled

2 for publishing. The audit raises concerns about

3 failing to classify these 20 data sets.

4   Second, there should not be confusion

5 about what a public—public data set is. The Open

6 Data Law defines a public data set a comprehensive

7 collection of interrelated data that is available for

8 inspection by the public in accordance with any

9 provision of law, and is maintained on a computer

10 system by or on behalf of an agency. In other words,

11 the data set is either wholly or partially a public

12 record subject—subject to disclosure under the State

13 Freedom of Information Law, or is already shared on

14 an agency website in another form. It is a public

15 data set, and should be on the Open Data Portal. The

16 Open Data Team has apparent confusion about the

17 definitions of the terms public, data and data set

18 are alarming, and the Administration needs to work

19 with Council and stakeholders to clarify and resolve

20 these definitional question or the Open Data Law--

21 Open Data Law cannot work. Thank you for the

22 opportunity to testify.

23   CHAIRPERSON VACCA: Yes, that's very

24 interesting. So, you're saying that they're not in

25 compliance.

2   DOMINIC MAURO: I'm saying based on the—

3 the results of the audit there are 20 data sets that

4 are not part of agency compliance plans, but are or

5 may be public data sets, and the audit is unclear

6 about whether these 20 data sets are, in fact, public

7 or not, but if these are public data sets, and

8 they're not part of the compliance plan, these

9 agencies are not in compliance with the Open Data

10 Law.

11   CHAIRPERSON VACCA: So some data sets,

12 nearly half of the Sanitation Departments are listed

13 as public, but not on the plan.

14   DOMINIC MAURO: I don't have the specific

15 number in front of me, but--

16   CHAIRPERSON VACCA: [interposing] They

17 are there.

18   DOMINIC MAURO: --there—there are four

19 things (sic) and yes.

20   CHAIRPERSON VACCA: Yes. Why would

21 something be listed as not public? Why—why is that?

22   DOMINIC MAURO: Well, as MODA testified,

23 it's not always clear where the line between one data

24 set and another data set is. So, they—they had a few

25 examples I think on the top of page 3--

2   CHAIRPERSON VACCA: Yes.

3   DOMINIC MAURO: --that are data sets that

4 may or may not be separate data sets or they may or

5 may not be public data sets, and those are important

6 questions, but we believe they should have been

7 resolved by this audit, and they remain unresolved by

8 the audit. I mean the—the Audit Report specifically

9 says that more—more investigation needs to be done--

10   CHAIRPERSON VACCA: Right,

11   DOMINIC MAURO: --but that was the

12 purpose of this—this audit.

13   CHAIRPERSON VACCA: And there's no answer

14 as to when the investigation will be done?

15   DOMINIC MAURO: Not in the audits. I

16 mean we—we hope to see MODA continue that

17 investigation with that. The report doesn't identify

18 this.

19   CHAIRPERSON VACCA: Well, let me ask if

20 you would like to come back. I'd like an answer to

21 this. Do you want to come back to the witness stand

22 please? You're already sworn in. Just identify

23 yourself for the record again?

24   DR. AMEN RA MASHARIKI: I'm Amen Ra

25 Mashariki, the Director of MODA. So in the report

2 that we released, we actually stated that 14 for

3 Sanitation, 3 for DOC and 3 for HPD, which make up

4 the 20 that Dominic is referencing, but we engaged

5 the agencies. We determine that there needed to be

6 further questions around the validity of those data

7 sets. We looked at them from the stance of MMR

8 metrics that went out, and so we said well what

9 datasets are you looking at that help you derive

10 those aggregate and then more of those (sic), 14 for

11 Sanitation, 3 for DOC, and 3 for HPD. Out of the 14,

12 12 of them are already on. We threw additional

13 engagement already on the Open Data Portal through

14 other datasets. Alright, so DSM (sic) reported that

15 for the following base sources associated with MMR

16 indicators, the most granular that the data can be

17 publicly represented according to the parameters of

18 Local Law 11, are the indicators themselves. And so,

19 we are not able to sort of pull those out into actual

20 data sets. Then there are two datasets, garbage

21 collection and road salt spread by borough that we—

22 DSNY has postponed a final decision on two pending

23 datasets nominations until the spring as they do not

24 have sufficient capacity to fully investigate until

25 after the snow season. And so they're pending

2 investigations. So we found that 12 at—at it's most

3 granular was what was put on the MMR metric site, and

4 then two, they're continuing to do work because they

5 don't have the resources. For Department of

6 Corrections RVM (sic) census contains mostly public

7 data, but the system is older and not suited for

8 automation. DOC currently meets the public need of

9 this data set through IIS data including the

10 currently published Daily Inmates in Custody data

11 set. IRS a large data set with conceptually dozens

12 of viable data sets. Some data is private, some

13 public. DOC is already publishing some of the public

14 data through the following data sets: Staff

15 injuries, Class A injuries, Inmate Incidents,

16 Slashing and Stabbbing and Inmate Incidents, Inmate

17 Fights and Inmate Assault on Staff and Inmate Deaths.

18 The third one from DOC is Programming and that MR.

19 The dataset is a decentralized non-systematized

20 static snapshot of high level aggregate statistics

21 not suitable for automation or publication. So

22 again, this goes back to the concept that the

23 investigation and the examination and verification as

24 we dug deeper into the datasets, we found that some

25 of these systems were set up such that it's very

2   complex and difficult to release, but a lot of it was

3   they found other ways to release it. Under HPD the

4   three data sets: Daily Updates with Charges data.

5   HPD is working through the requirements and

6   developing a road map to transition to daily updates

7   of all code enforcement data, AEP data. HPD is

8   putting together requirements for the AEP data. They

9   need to put this data into the Business Intelligence

10   Systems, and put together the documentation. Their

11   plans for the Code Enforcement files are the

12   priority, and then Housing Production data files for

13   NHNP. HPD does not plan to release the NHMP data

14   because of data quality concern.

15   CHAIRPERSON VACCA: Is there a deadline

16   that you give these agencies where you see situations

17   like that that they must post that information that

18   the—when is the investigation going to be complete

19   and when, you know. One by one do you audit and then

20   tell them what you expect in so much as the

21   outstanding sets?

22   DR. AMEN RA MASHARIKI: We work with

23   them. We don't give them a deadline because again

24   with the limited resources we have, we can sort of

25   begin to require their IT people. We don't know what

2 their IT infrastructure looks like, and we can't

3 begin to, and so but we do engage with them to get as

4 quickly and prompt a feedback as possible. As you

5 can see, where we saw the 14, 3 and 3, the 20 data

6 sets, when we saw those as possible public data sets,

7 we engaged each agency individually and—and walked

8 with them through a deeper dive and we pulled out all

9 of this information. And so, instead of just saying

10 hey we want an additional answer within the next few

11 days or week or say, we say, hey, how can we be

12 helpful in engaging with you to get to the bottom of

13 this. So that's the strategy we have thus far.

14   CHAIRPERSON VACCA: Well, I will tell you

15 that a concern I had legislatively when we went ahead

16 with much of the legislation was how does one city

17 agency tell another city agency what to do? How do

18 you force them to do something? That was why I

19 involved DOI basically as the examination—

20   DR. AMEN RA MASHARIKI: [interposing]

21 Yes.

22   CHAIRPERSON VACCA: --in the examination

23 process. So you seem to be articulating that concern

24 yet today that you have to cajole to get compliance,

25 and that worries me because the time table seems to

be open-ended on what's outstanding, and I think that

that should be addressed. We just can't have open-

ended-- You know, we have to make a decision

basically what is compliance and what is not

compliance. That—that is the decision we have to

make. Either you comply—the fact of the matter is

you comply or you don't comply. There's no

alternative fact. I'll put it that way. So, hey,

you don't comply. So I—I don't get it.

DR. AMEN RA MASHARIKI: Yeah.

CHAIRPERSON VACCA: Certain—certain

things I'm hearing are—certain things I'm hearing now

is that you do know step-by-step what's going on,

which I appreciate, but we have an open-ended

compliance situation so we--

DR. AMEN RA MASHARIKI: [interposing] So,

one, I think you used the term cajole. I think the-

the term I used was engage, and we have a great

relationship with the agencies, and no agency has

pushed back, and it's been a great back and forth

with the agencies. Two is one of the key terms of

the law is examination, and this—as this being the

first, it allowed for us to jump in with both feet

and understand the landscape that was planned, and so

2   it—it-it—this affords us an opportunity to go back

3   and rethink our strategy and what—which we've said

4   this has always been a continuous improvement

5   process. And so, yes, we did find some things that

6   we did not expect to find, but my sense is that's

7   exactly what you get from an examination is you get

8   your sort of the state of affairs. So we got the

9   state of affairs, and we're working with the agencies

10   with what they have in terms of resources, and

11   capabilities to move the ball. But the Open Data

12   Report will always be the final statement on who's in

13   compliance and who's not in compliance. That's what

14   we will always use to state who's in compliance—who's

15   compliant and who's not compliant. So you see

16   anything out of this examination and verification and

17   review, if there is non-compliance, you'll see that

18   in an Open Data report. That will always reflect

19   that.

20   CHAIRPERSON VACCA: You mentioned

21   resources. If an agency does not have resources,

22   who's requesting the resources from the

23   Administration?

24   DR. AMEN RA MASHARIKI: It would be the

25   agencies.

2   CHAIRPERSON VACCA: The agencies? In

3   your opinion, do agencies—in your opinion, are there

4   now agencies that don't have the resources to

5   totally—totally comply with this legislation.

6   DR. AMEN RA MASHARIKI: Oh, yeah. We've

7   done Examination and Verification of three agencies.

8   So I can speak to those three agencies and they do

9   have the resources.

10   CHAIRPERSON VACCA: Moving forward, we're

11   going to be getting to smaller agencies at a certain

12   point. So I want you to keep that in mind.

13   DR. AMEN RA MASHARIKI: Absolutely.

14   CHAIRPERSON VACCA: And if—if it's a

15   resource issue, then you would go to the

16   Administration and say there's a resource issue here

17   when complying with the law?

18   DR. AMEN RA MASHARIKI: We would engage.

19   CHAIRPERSON VACCA: You would engage.

20   DR. AMEN RA MASHARIKI: Yeah, and—and--

21   CHAIRPERSON VACCA: [interposing] At that

22   point in the budget process?

23   DR. AMEN RA MASHARIKI: That's right.

24   MODA wouldn't (sic) do it.

25

CHAIRPERSON VACCA: Okay. Now, I wanted

to ask Dominic do you recommend any changes? Does

your organization recommend any changes to Local Law

8, any—is there any legislation you foresee?

DOMINIC MAURO: Not to look at it no. I

think that the—that the time table for the—the, you

know, it's implementation year. I'm happy to hear

that this is going to be part of the—join with the

Compliance Report in July. It makes a lot of sense,

but no that's specific recommendations.

CHAIRPERSON VACCA: Okay.

DOMINIC MAURO: To the Local Law.

CHAIRPERSON VACCA: Okay, any questions

from the members. No. Okay. Well, I want to thank

you both, and appreciate your testimony as well, and

we—we look forward to working on this diligently with

you in the days ahead and we appreciate your hard

work, and your-your outlook for future, which I know

will be even more—will lend to even—will lend to even

more transparency in government.

DOMINIC MAURO: Thank you.

CHAIRPERSON VACCA: Without other ado, it

is now almost 11 o'clock and this hearing of the

1   COMMITTEE ON TECHNOLOGY   49

2   Technology Committee of the City Council is

3   adjourned.   [gavel]

4

5

6

7

8

9

10

11

12

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   February 2, 2017