

CITY COUNCIL  
CITY OF NEW YORK

----- X

TRANSCRIPT OF THE MINUTES

Of the

COMMITTEE ON TECHNOLOGY

----- X

December 6, 2017

Start: 1:02 p.m.

Recess: 2:10 p.m.

HELD AT: 250 Broadway - Committee Rm.  
16<sup>th</sup> 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, Mayor's Office on Data Analytics

Craig Campbell, Special Advisor  
Mayor's Office on Data Analytics, MODA

Albert Webber  
Dept. of Information Technology & Telecommunications

Noel Hidalgo, Beta New York City

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



[sound check, pause]

CHAIRPERSON VACCA: Everyone please take your seat. I thank you all for your cooperation. As you know, I like to start on time because all of us have time that's valuable, and I treasure your time, and I thank you for treasuring mine. So, we do our best to be prompt. I do want to note I'm joined by our newest Council Member Adrienne Adams and we welcome you to our meeting today. This will be my final meeting of the Technology Committee as chair, and I want to thank all of you. Many of you have come every month, many of you have worked with us on issues and problems and legislation and policy, and it's been a pleasure to chair this committee. We are moving forth with a very important legislative piece on Monday. No, tomorrow. Tomorrow at 10:30 we have another committee meeting for a vote, and I think that originally when I was made Chair of the Technology Committee I didn't think that we would accomplish all we have. It really goes to the point that a Chair of the Committee makes the Committee. A chair can do so much even with a committee that doesn't sound like it has much meat and potatoes. This committee is not Land Use. It's not Finance,

but boy, we did a lot, and all of you gave fantastic input in the legislation that we put together. I know I'm going out of sequence, but I do want to thank Malika and Patrick for their fantastic work with me, and I want to thank my Legislative Director Zach Hecht because without him, forget about it, as they say. A member is only as good as his staff, and members depend on their staff's expertise. A good member depends on his staff to tell them when he's crazy and, of course, they want to—you want to hear when they're good, but I would rather my staff not always tell me how good I am. I'd rather that they tell me when I'm wrong because that's what they're there for. I don't want them to bow to me. I want them to tell me Jimmy, are you out of your mind? What are you talking about? And then we arrive at what we have to do. So, enough of that, but I thank you again. So, now to the script. Good morning, everybody.

COUNCIL MEMBER: Good morning.

CHAIRPERSON VACCA: It's really good afternoon. First of all, whoever wrote this made a mistake. This is a terminal mistake. This is the afternoon. What am I going to do? I just—

MALIKA: Don't blame it on me.

CHAIRPERSON VACCA: Malika, I just played you as being wonderful. [laughter] What am I going to do? Let me correct my Counsel. Good afternoon. My name is James Vacca, and I'm Chair of the Committee on Technology, and we're here today to discuss three agencies' compliance with the Open Data Law, and to discuss the Mayor's Office on Data Analytics, examination and verification 2017 Findings Report. Open useful data that can be shared with the public is a cornerstone of a transparent government. Data can help New York City residents improve their neighborhoods, advance research, improve public safety and provide countless other benefits. We passed the City's first Open Data Law in 2012 requiring eligible city data to be published on a single web portal by 2018. We have continued to improve the law in subsequent years through a series of amendments. One of those amendments is Local Law 8, which is the subject of today's hearing. Local Law 8 requires MODA to create a plan to assess how select agencies are comply with the Open Data Law, and to uncover data sets that might not have been disclosed when agencies self-report. The agencies

undergoing MODA's examination this past year, include the Department of Buildings, Fire Department and the Department of Environmental Protection. The results of the report do raise questions about some agencies' commitment to Local Law 8, and the resources MODA may need to assure accurate reporting. For instance, it appeared that some work sheets in the report were blank, and there was no explanations. Our committee staff also noticed gaps in the data that was reported in the worksheets. We commend DOITT and MODA for helping New York City be a leader in making public data available at the municipal level. Our City government creates massive amounts of data everyday, working with all these agencies to ensure their compliance with the Open Data Law, managing the web of formats agencies may have traditionally used to publish data, and coordinating the publication of this data on a single open data portal is no small feat, and MODA has already done a lot of great work. We are here today to ensure that the city's open data continues to improve and that we make strides to assure New York City is a leader in open government. We want to get additional insight from MODA on its 2017 report. We also welcome advocates who will

testify today on ways to strengthen city agency compliance with open data, and again we thank you for your help. I help the Administration. I did not thank the Administration before. I thank them now. We've had excellent cooperation from the Administration on the Technology Committee all along. It does appear to me we have more to do, and this oversight hearing is designed to see where we have gaps and what we can do about that. So, we do have three witnesses here. I'd like you to please engage me in the affirmation, which the Chairs administer. Please raise your right hand. Do you swear or affirm to tell the truth, and to answer truthfully to the Council questions that they may pose as we proceed?

JAMES PERAZZO: [off mic] Yes, we do.

CHAIRPERSON VACCA: Okay. Thank you.

Now, would you—who would like to go first, and would the first person please introduce themselves for the record?

JAMES PERAZZO: Good afternoon. My name is James Perazzo, Deputy Director for Strategic Management of the Mayor's Office of Operations and Acting Director of the Mayor's Office on Data Analytics, also known as MODA. I'm joined today by



Craig Campbell, Special Advisor at MODA and Albert Webber from the Department of Information Technology and Telecommunications. On behalf of the Administration, I would like to thank the Committee on Technology and Chair Vacca for their unwavering support of the NYC Open Data Initiative. The Committee and the Administration are united in a commitment to transparency in government. The Open Data Law brings this commitment into community organizations, classrooms, and digital laboratories all across the five boroughs. New Yorkers mine and manipulate our data to develop solutions to a wide variety of issues while sharing information on how city programs operate, and what they (coughs) achieve. The open data policy that was born from the law is unparalleled among American municipalities and the amendments this committee has passed over the last three years will ensure that the program thrives in perpetuity. One of (coughs) one of those amendments is Local Law 8 of 2016, the Open Data Examination and Verification or E&V. This law requires MODA to work with three agencies each year to verify their compliance with the Open Data Law. Following the results of MODA's first examination and

verification report, which we discussed before this committee in January of this year. MODA developed a more robust examination process for the 2017 cycle. On December 1<sup>st</sup>, we submitted our findings on the Department of Buildings' Environmental Protection and the Fire Department in a report that is publicly available on the Open Data website. This—the process was a success. Each agency identified new data sets for publication and new ways to improve data sets already on the Open Data Portal. Further, the process affirmed each agency's commitment to open data. While the full examination and verification process is described in detail in the report, I will briefly highlight its major elements here. The first part of the examination was a data set questionnaire, which required agency Open Data Coordinators or ODCs to conduct—to conduct a thorough inventory of data sets at their agency. This consisted of two main components. The first was an inventory of the information technology and data systems that support the agency's core business processes. This required ODC's to describe each system, its users, and its supported operations to ensure that the agency's mission critical data sets are accounted for in the

examination. The second was an inventory of instances where data is already formatted and transferred at the agency. This required ODCs to identify the existing channels where public and internal stakeholders can assume the agency's data. Such as operational interagency data exchanges, public reports supported by agency data sets, data provided for the Mayor's Management Report and data used to fulfill Freedom of Information Law requests. Finally, the agency commissioner, or their designee was required to certify the completeness and accuracy of the inventory to ensure the full participation of relevant staff and the process. In addition, the agency general counsel or their designee, was required to sign off on the data inventory certifying that the process was conducted in accordance with statutory parameters of the Open Data Law. In completing this process, each agency identified at least one new data set to publish on the Open Data Portal. DEP identified data sets on the quality of water in city reservoirs and the geographic location of catch basins. DOB identified its certificate of occupancy dataset, which the agency published online in June, and next month the Fire Department will

publish the location of fire alarm—fire alarm call boxes, and an inspections data set both identified during the E&V process. The examination also included a survey of data quality for data sets already on the portal. Publishing data online is only the beginning of making data open. Truly opening data requires updating it regularly, identifying and fixing errors, and engaging users and incorporating their feedback. All three agencies identified ways to improve their existing datasets on the Open Data Catalogue. DOB (coughs) made data quality improvements to an existing dataset while DEP and FDNY identified datasets that should be archived from the portal. Following the examination process, MODA met with open data coordinators at the surveyed agencies to discuss the challenges and opportunities they face in their role. Insights from these discussions are reflected in the recommendations for Better Citywide Compliance section of the report, which lists ten specific actions for MODA, DOITT and city agencies to take (coughs) to improve the Open Data program. For your convenience, these recommendations are appended to today's testimony. These recommendations fall into two categories. The

first pertaining to MODA and DOITT, which should continue to provide strong leadership and proactive support to agencies in order to affirm the administration's commitment to the success of open data in every area of government. The second set pertains to agencies, which should leverage the assistance of MODA and DOITT in embedding open data into existing processes. To be successful, open data must not be a one-off reporting requirement that are a foundational element of data analytics and data sharing within city agencies. [background comments]

Given the pace technological and administrative innovation as well as personnel turnover and organizational shifts, identifying all public data sets at a given agency is a constantly moving target. The discovery of new data sets through this year's E&V Process does not mark delinquency on the part of any open data coordinator. Rather, it highlights the need for routine assessment to ensure that all business areas, data systems and communications channels are synchronized with open data publishing requirements. These insights are helping MODA build additional mechanisms to improve citywide compliance with the Open Data Law. For the annual report

published last summer, all agencies were required to submit an abridged E&V dataset questionnaire. This year, we will look to make additional improvements to the reporting process for all city agencies. The strength of the New York City Open Data Program is that it touches every city agency and engages hundreds of thousands of New Yorkers each year. Our ongoing partnership with the Open Data Coordinators and other Open Data Stewards at participating agencies is helping MODA build a more transparent city government. We wish to thank them for their full cooperation in this process, and for continuing the good work to identify, prepare and publish valuable datasets through the Open Data Portal. Finally, before concluding, I would like to state for the record that the city of New York is very lucky to have such a talented, hardworking and dedicated open data team. The Committee is aware I have been serving in the interim capacity at MODA for just a few months. Within that time, I have been deeply impressed by the expertise and ethics of this group, and I know New Yorkers are better off for having them on the team. At this

point, we welcome the committee's questions on the Examination and Verification process and results.

CHAIRPERSON VACCA: Thank you for your testimony. I have some questions. Let me ask one or two right off the bat, though. How many people on a daily basis use open data? I'd like to know how widely used it is?

CRAIG CAMPBELL: The last report that we pulled, which was yesterday--

CHAIRPERSON VACCA: [interposing] State your name.

CRAIG CAMPBELL: Craig Campbell, Special Advisor at MODA. In the last report that we pulled we had 155,000 monthly sessions. So, if you break that down into daily, we're looking at about 5,000 daily users.

CHAIRPERSON VACCA: I would like to know, as--as an exercise in comparing and contrasting, what other city agencies get by way of hits. Where do we rate? I--I would think that we rate pretty low. Even that number surprises me, to be honest. I just question if we've done enough to get people to use open data to know how to use it. I--I just have spoken to a lot of people about it, and they don't

seem to be too much aware of it. It's a tool that's fantastic, and you and I can agree to that, but unless we engage the public more, then the tool has not become as successful as the—as it should be, as we envisioned in the original legislation. So, my question goes to engagement of the public outreach. Are you at the community boards? Are you at community associations? Are you at churches? I—I'd like to know what are we doing with engagement of the public?

JAMES PERAZZO: Uh-hm. Well, I would like to (coughs) thank our—our friends in the advocate community and in the civic tech community for moving forward with a lot of these efforts already. Noel is here and will testify later, but has done a terrific job working with the Manhattan Borough President's Office to engage community boards. In addition, we have identified ways, existing touchpoints where people interface with their city government and democratic processes. One example of this is the participatory budgeting process. We developed resources for them to distribute to every one of the 70,000 New Yorkers who are engaged in that process so that the—their



proposals or projects could be informed by city data. We are additionally looking into similar partnerships with the libraries and—and we'll continue to search for opportunities.

CHAIRPERSON VACCA: We need someone from the city to do the same engagement that I'm sure Noel has done in Manhattan in the Bronx and Queens and the other boroughs. I have community board members in the Bronx who do not know about open data, and we do not have the training. I know Gale Brewer has made it a priority in Manhattan, and she's a—she was chair of this committee years ago, and this is her—her thing, but we need you to make the other community boards more knowledgeable about open data because they are the ones who can use this information as they monitor city services on a day-to-day basis. Of course, Council people. No doubt Council people, but when you do the community boards, you can involve City Council people offices there if they want to staff to the community board training. So, what are we doing in the other boroughs in so much as training board members? This may have to come from your office because I don't think that we have the—the linkages that Manhattan has.

CRAIG CAMPBELL: Uh-hm, uh-hm. Using lessons learned from preliminary trainings that we've had, working with Beta NYC, working with the Parks Department and their Computer Research Centers. We've had trainings in Manhattan as well as the Bronx. We've done a community board training in Brooklyn. I think a lot of these efforts can and should scale, but we have been looking at all those different touch points.

CHAIRPERSON VACCA: That's a very vague answer, though. You should be in politics. I need specific answer. What are we doing in-in-in-in Brooklyn, and the Bronx rather and Councilwoman Adrienne Adams is from Queens. So I'm sure she's interested in our boroughs. Where are we in these two boroughs with open data training? I have to tell you as a Bronx councilman I know that we're not anywhere. So, what's the next step? What can we-- what can we look forward to?

JAMES PERAZZO: One specific thing that we can look forward to is coming up this March will be the second ever New York City Open Data Week, which we'll be producing in partnership with our friends in the advocate community and communities.

That effort has—last year started about a month before the anniversary of the Open Data Law. This year we started far more in advance and plan to—we hope to have events in every borough, scale up to the level of engagement and the 900 users that we engaged in the first iteration of this week, this year and continue to again identify, you know, that great list of—of partners and—and resources.

CHAIRPERSON VACCA: Yeah, I'll end it here. I mean my suggestion is March is fine. You're going to do something in March, I would partner up with City University, use the colleges. They have large computer labs. Try to get community boards as your population to come in, members' offices, district managers they should know, and I wonder how many of them know about how to use open data. But if you use—if you use CUNY, Queens has several colleges, the Bronx has several colleges, and that could be a point where you could have training. Publicize it, get the members, give them advanced notice of it, but I—I would think that that may work, and that's my suggestion to you. Okay, Barry Grodenchik was here. He had to leave. I want to mention that. How is the

Open Data and examination and verification process  
beneficial to agency Open Data Coordinators?

JAMES PERAZZO: The process gave the  
appointed Open Data Coordinator an opportunity to  
review the entire organizational chart, and locate  
each of the lines of business where there is data  
moving from a source system. So, it required a very  
thorough process in addition to the annual reporting  
requirements that we asked them to report on in the  
report that we published on July 15. It also  
required them to work with all of the different, what  
we termed in the report Public Information Stewards  
at the agency. So, the Open Data Coordinators were  
working with their CIOs, their CTOs in some cases and  
leadership as well as public records officers and  
general counsel.

CHAIRPERSON VACCA: Now, in the  
Examination Report, you mentioned in one of the  
comments that data sets are loose. What does it mean  
when you say that? MODA says that some of the data  
is loose. What does that mean?

CRAIG CAMPBELL: A lot of city data as it  
rests in the source systems can be queried and-and  
represented as different kinds of tables at different

levels of granularity, and so the same fundamental information can be rolled up into an array of different data sets, and that's what we meant by the term.

CHAIRPERSON VACCA: [off mic] And now when a—when an [on mic] Open Data Coordinator leaves, do you have a transition process? Do you train the next person? What do you do with that transition period? Then, do you appoint someone new within that agency? How do you handle that?

CRAIG CAMPBELL: Uh-hm. So, that was one of the recommendations that we specifically called out in the report that we can do a better job with. We also listed in an appendix the current engagements that we've had with Open Data Coordinators over the past year, some of which include doing that proactive transition support, but through our survey of Open Data Coordinators and our deep dive interviews with the agencies that were examined, we identified a need for additional resources for Open Data Coordinator turnover.

CHAIRPERSON VACCA: Now, we noticed one thing. When our staff reviewed this, they noticed about verification. We observed that the Fire

Department did not answer any of the questions in one of the worksheets, Worksheet D, even though they have data that should have been included. So, for example when they were reporting on firefighting response times to emergencies, and then they were reporting on smoke detectors and alarms in fire-related deaths, those deaths were not reported. How do you verify the reports that agencies submit to you? How do you make sure that the reports are all inclusive as per the mandate of the legislation?

CRAIG CAMPBELL: So, our process is described in the report. We have the agencies do this thorough review, and we work with all of the relevant stakeholders at their agency, and we return us the workbook. We then looked at the workbook, identified gaps, additional questions that we had and returned it to the agency, which went through another round of review again with all of the relevant stakeholder including their Legal Affairs Offices. For the Fire Department what they reported in that worksheet was what they certified as accurate.

CHAIRPERSON VACCA: [off mic] In this case it wasn't—[on mic] In this case it was not all-

inclusive in the case of the Fire Department. Am I right?

CRAIG CAMPBELL: If there is a specific item that you think was missing then we'd be happy to take it back to the Fire Department.

CHAIRPERSON VACCA: It was—it was the ones in the question. Those—those were cases where we identified that the Fire Department did not provide the all-inclusive information that was required. So, then what is the next step?

CRAIG CAMPBELL: The purpose of—what we view as the purpose of this report and all of our public reports is that it's a process of continuous improvement, and if there is a specific area that in this certified workbook that public users believe is missing, then that's part of the process. Examination and Verification doesn't end on December 1<sup>st</sup> when we submit the workbook to you. It's part of an ongoing process with each of these agencies.

CHAIRPERSON VACCA: But there are deadlines that have to be met. So that in this case the Fire Department did not meet the deadline. The information is not there. We count on agencies to meet the deadline. That's why we have legislation

that prescribes those deadlines. So, we now count on someone to make sure that the legislation is adhered to. So, who does that? Is that your office or do you just say to the agencies we count on you? There has to be someone that is assured of accountability. That's my point.

CRAIG CAMPBELL: And we agree. Again, what the Fire Department reported to us was what they certified as accurate. There are specific inaccuracies that the Council Member and other public users notice we'll take back to the Fire Department.

CHAIRPERSON VACCA: Do you need more resources to assure accuracy or in this case does the Fire Department need more resources to assure accuracy? Where—where do we—where was the ball dropped here or was the ball dropped? I just need to know that.

CRAIG CAMPBELL: [pause] We—I think additional resources are always helpful. It was cited by the Open Data Coordinators in the follow-ups that we did with them that—that one of the challenges they face with the Open Data Law is that the Open Data Law does not furnish the resources to achieve all of the goals, and it is often a requirement added



to an additional role, or adding to an existing role that they have within the agency. So, I think always additional resources for Open Data Coordinators and for our team are beneficial, but I also stand by the report that was noted.

CHAIRPERSON VACCA: Alright, if we say that additional resources are needed, the Mayor is releasing a November budget. A November plan is coming out I think Monday. Has your agency worked with the Mayor's Office in identifying possible additions to the budget that you would request in the expense allocations that he's going to be putting forth in the—in upcoming budgets especially in January. That's when the Preliminary Executive Budget is released for the next fiscal year. Have you requested help that you think you may need?

CRAIG CAMPBELL: We—we have not.

CHAIRPERSON VACCA: Well, now you have to. We have to do that. You're telling me publicly at a hearing today that you need more resources. Then there has to be a submission that you prepare in my opinion since you've gone on the record saying you need more resources. The next step is to advise the Mayor's Office through your agency what are the

resources you need in order to complete and carry out the charter mandated functions. Am I correct?

JAMES PERAZZO: Well, as--as Craig said, we--we believe that those functions are being carried out, and to the point about specific, what may be specific items that are missing, we can take that back to the Fire Department and investigate the issue and--

CHAIRPERSON VACCA: It seems to me that you needed to show resources. It was stated for the record and from what I can see when I spoke about community outreach, I do think that there's a need for resources. We've--we've found this discrepancy here, which was not compliant. I know that you many not be in a position to commit to doing that today. That may have to go higher up in your agency, but I would certainly suggest as a result of this oversight hearing that this be taken back--taken back, and that you start discussing this in so much as it pertains to Preliminary Budget--Preliminary Executive Budget in January. Okay. So, that's my recommendation to you. I do understand your position is such that you may not be able to give me that assurance, but I do think that it's become evident that that should be the

case. I certainly want this program this program to succeed as a major innovation that the City of New York has undertaken, which is ahead-ahead of the rest of the nation, but in order for it to succeed, there have to be resources in place. Now, I saw something that DEP was releasing two data sets, and they- they're going to release them on December 31, 2018. Why was that date chosen, December 31, 2018? I don't think many people are going to go on open data on New Year's Eve. I don't know. Tell me is-is it-is that-is this-can we get that out before that?

CRAIG CAMPBELL: DEP has assured us that they will take every effort to publish those datasets before that date. I think that the one that they submitted in their plan was cautious given that those datasets can only be published as a result of a technology upgrade that is happening next year.

CHAIRPERSON VACCA: I appreciate that. I was thinking of Washington for a second because bad news always comes on Fridays because, you know, they think the people on Saturday are sleeping and, you know, that they want to be left alone. So, on New Year's Eve, only someone like myself will be on the Internet on New Year's Eve looking fore the open data

sets, you know what I mean? So, I want to make sure that you understand that we want maximum public participation. Have been--this is a good question. We were talking at another hearing about FOIL, and I wanted to know if open data has reduced FOIL requests?

CRAIG CAMPBELL: We have anecdotal evidence from the agencies that we work with that publishing data has reduced FOIL requests.

CHAIRPERSON VACCA: Alright, you have anecdotal information. I think that we need specific information. Can I ask that you do a worksheet for the committee indicating specifically if that's happened, and to what degree? I'd like to know to what degree. I think it's helpful to city government that we have an open data system that may end up reducing FOIL requests. People complain about FOIL requests taking weeks and months to have--to get answers back, but if it's on open data, you just go online and get it. So, it's certainly better to get it under open data, but I would like to know the degree to which that has occurred. Okay, and get back to the committee?

JAMES PERAZZO: Certainly.

CHAIRPERSON VACCA: Thank you. Before I leave office December 31<sup>st</sup> because I don't have time this here, you know. Okay, anything else? Does Council Member Adams? No, no questions? Council Member Adams.

COUNCIL MEMBER ADAMS: This is just to really echo Chair Vacca, and my concern of the information not really reaching places that it really, really needs to reach. He emphasized community boards, and I can let you know that this is an extreme resource that we're--and an opportunity that you're missing out on that really, really needs to be looked into, and taken advantage of post-haste in your own behalf. So, I just wanted to echo the sentiments of Chair Vacca.

CHAIRPERSON VACCA: Thank you, and I want you to know that Council Member Adams and I come here with a expertise on community boards. Council Member Adams is a former Chair of a community board. I'm a former District Manager. So, we've been grassroots people, and we know that this is something we love o our members to have, but they don't have that skillset at this point. Okay. I have no further questions. You want to say anything else?

CRAIG CAMPBELL: No, thank you.

CHAIRPERSON VACCA: You're just glad I have no further questions I think. That's what happens here. Alright, thank you, and we will call our first panel. Noel Hidalgo has been one of our fantastic advocates. John Kaehny, Reinvent Albany, Noel is with Beta New York City. [background comments, pause] Can I—can I just ask for a one-minute adjournment. I just want to use the men's room. I'll be right back.

NOEL HIDALGO: Sure. [pause]

CHAIRPERSON VACCA: Yes, we'll reconvene. I apologize. Noel, would you want to start? Oh, John, you want to start? Introduce yourself for the record, please.

JOHN KAEHNY: Sure. My name is John Kaehny. I'm Executive Director of Reinvent Albany and Co-Chair of the New York City Transparency Working Group. Thanks for having me today, and—and thank very, very much again for all the work you've done for the Open Data Initiative here in New York City. I think that you've—you've carried us forward wonderfully and gotten us through some bit of a rocky time during the transition, and it's your leadership

has really been crucial, and I agree that a chair can make all the difference. So, thank you so much for all your efforts. It's really, really appreciated. So, thanks.

CHAIRPERSON VACCA: Thank you and I'm not going to give you a time limit today because you've been so nice.

JOHN KAEHNY: Well, thank you. The hearing today is, of course, about the-the MODA audits of the agencies, but it also has a larger purpose that you spoke to in your remarks, which is to keep putting-keep pushing the city forward on open data, and by focusing on specific agencies and finding systemic problems generally from those specific audits. But let me just speak for a second to the larger purpose. The city's Open Data Program did have a bit of a rocky time during the transition between the Bloomberg and de Blasio Administrations, and I wanted to flag something right now. We're again concerned about staffing and resources for open data, which in your questioning you touched on, in fact. Specifically, seven months ago, Rafael Corvalho, who is one of the two Senior Open Data Managers for DOITT, left to go to DCAS, and he is a

very talented manager. He did important work. Unfortunately, as far as we can tell, he has not been replaced, and DOITT does not have a hiring process going on and has shown either an inability or a disinterest in replacing this important position, and I'm mentioning this at the onset of my testimony for a reason. DOITT's folks, their senior managers are the key lynchpins or the key multipliers that make open data work, and they have to keep on pushing and pushing, and we feel that the lack of a person in this key position in DOITT has slowed and undermined the citywide open data effort. So, we're—we're concerned about that, and I'm mentioning that because cause we would like you in your position as Chair to ask DOITT to replace that position, and particularly in light of the fact that it's a funded position as far as we understand it, and that there my be cuts coming up and that job needs to be filled. And we're flagging this as a big problem right now because seven months missing a key open data manager is a problem, and that—that position has got to be filled or we're going to see continued lack of progress in some key areas in open data



CHAIRPERSON VACCA: I will. I will reach out and try to work collaboratively with DOITT, and to see what has happened there. I think that you raise an interesting question, which the next Council has to get a handle on more than this Council is what has been the impact of the "job freeze" quote/unquote? We are under some type of a job freeze. I'm not clear on the scope it. Where many people thought it would be temporary in nature we don't know if that's the case. We don't know what titles have been affected in what agencies, and what it's meant for services that we expect agencies to provide. So, I do think you're raising a significant policy issue that has to be addressed, but separate from that, I will reach out to DOITT, and we will have a conversation.

JOHN KAEHNY: Great. Thank you for that.

Now onto the audits and some comments about the audits, and we think the law is very valuable. In fact, more valuable as it progresses than we even thought when you first introduced it, and it-it-it-both because it asks MODA to directly engage some of the larger agencies and push them specifically in a way that they normally wouldn't, but also because

MODA has used it effectively as their testimony suggested to help inform their bigger efforts, and we think it's great that they're using the Open Data Audit Questionnaire in a—in a lesser—in a smaller form, summary form for the Annual Open Data Reporting, which is a—a—just a smart thing to do, and that's exactly what they should be doing with these audits is finding ways of continuous improvement, which we agree that open data needs from the—the more intense drill-down that they do in this audit. So, compliments on that to both you and MODA because this is showing that this law is broadly working in that way. So, that's—that's a great thing. That said, we think this law is not nearly as useful as it could be to the public and to Council. Specifically, we'd like to see a concise narrative or summary that just simply comes out and says how well an agency is doing because right now it's easy to miss the forest for the open data trees here. I mean even for people like us who do a lot of open data policy and system analysis, it's just hard to tell from the audit whether or not an agency is a leader or a laggard, and we would like to see MODA just come out and simply say this agency is doing a great job at X, Y,

Z. They might need some work on these other things and here's how know this from the—the questionnaire and the interviews and the public indicators that we've provided with you. Because what we have now is a series of spreadsheets with many, many entries in them that we have to spend a ton of time creating context around so that we can understand them. We have to go back and look at the Open Data Portal, and what that agency has posted there, and look at the—the kind of data that they have up there, and the use around them to create a useful context. And the reality is that, you know, no one is going to do that on the public side, and so right now this lacks just the simple story around those—those data points, and those questionnaires. You know, we simply want to know MODA tell us, you know, is this agency a—a great performer or are they doing very badly, and this audit does not tell us that right now. We have to read between the lines and we have to provide a lot of context for that. So, we would urge, if possible, that there be some kind of commentary on it. I mean MODA knows what's going on, and they're not really saying it in these, and the politics can be difficult. But you foresaw this when you originally

asked the Department of Investigation to do this because you thought there were a more independent agency that could speak the truth on this, and we do feel that it's a, you know, it subtracts from this overall effort and work that MODA just can't say here's what's going on here. So, that's—that's my biggest comment. That said, from looking at this information, we think that the Department of Buildings is clearly doing a great job, and kudos to them because the audit and the completeness of their responses and the many comments they have show that they're really engaging with—internally within their agency about open data, and also with their stakeholders and the public. I mean it looks like a vital active thing that's going on at DOB, and I would contrast that with the Fire Department, and I hope I'm not wrong, but based on the information provided in the audit, they seem to be doing a bad job on open data. Now, again, I don't know. I wasn't there with the audits, and without, you know, maybe MODA would dispute that or the Fire Department, but what the questionnaires suggest and the—and the information provided to the public, and the Fire Department as you noted, did not fill out the entire

important worksheet. It's just simply not there, which is the one that says: Which are—which datasets are you using to provide reports—mandated reports under city law? They simply didn't fill that out, which—which is pretty poor, and I, you know, I wish MODA would have not let them do that. But I would say that out of those audits, out of the first six agencies looked at, the Fire Department looks like they have a real problem here, and that's out of again, you know, maybe that's no right. Maybe it's just someone who hastily filled out the worksheets or whatever, but of the six agencies that we've seen so far under these two years of audits, they jump out of the worst performer, you know. So, that's—that's what we're seeing right now, and so we would like to see them respond to that and fill out worksheet D, and maybe provide an explanation of how they're getting there. Now, going forward there are some things missing from the audit, too, in terms of just telling the simple story of how they're going to meet their deadlines and mandates. It just does not come through in the audit. So, Council has had this process of continuous improvement in what you could think of as agile legislation, which is legislating

as we learn new things, and the audits don't tell us how the agencies are going to get to their data dictionary or geospatial standardization targets. They just have a list and column. There's no--there's no real information that lets us know if they're going to get there on time or not, and going back to the Fire Department, I'd not that I think they had 17 datasets and only a handful had data dictionaries and the rest I believe three and the rest of those aging--data sets were going to be done by your favorite date December 31<sup>st</sup>, but in this case of 2017. So, how are they going to get there in a few weeks on data dictionaries when they haven't been able to bet there in months? You know, I don't know, but that's what the audit says. So, some commentary that says here's how we're meeting our plan. You know, here's what we're focusing on. Here's what we think and--and both MODA and the agency talking about they think they do well, and what they do wrong. That would be invaluable insight because otherwise we just have these columns and worksheets, and we could misinterpret them very, very easily. I have a list of question I'm going to share with you, and with DOITT as well, and I'm glad you asked the question

about what does it mean when you have loose dataset?

We mentioned this last year in our testimony about

the open data audit, but we still are a little

concerned about whether or not the agencies

understand the Open Data Law. Hopefully, they do,

but it's not clear to us that they understand what a

public record is completely based on what datasets

that they're identifying in their publishing

schedules. So, I'm just going to flag that there as

something that needs more education especially in

agencies that don't have a lot of public

stakeholders, and really pushing on them like DOT or

DOB might have. There are some questions that are

just missing that the audit could easily include,

including any stories about-with-from agencies about

having open data making their life easier, and one of

the complaints before we passed the Open Data Law was

that people within agencies had trouble retrieving

data from their own agency that it was siloed or

difficult to get to, access or understand. So, then

it would be invaluable to hear about whether the Open

Data Law has made it easier for people within

agencies, and from other agencies to access that

agency data. That's something we just can't tell. We

don't know and we can't see from the audit. So, that would be helpful, too, and lastly, it would be very hopeful for the audit to include interviews, short conversations from agency stakeholders because we haven't heard anything at all from the public, and they often know exactly what—how well the agencies are—are doing and not doing and particularly big agencies like DOB or agencies with a lot of public stakeholders like City Planning that use a lot of open data. So, anyway, thanks again for all that you've done for us and I'm glad—glad to be here today to mark this day as your last day with Technology. So, thank you.

CHAIRPERSON VACCA: We've been joined by Council Member Greenfield. Noel.

NOEL HIDALGO: Good afternoon. I want to say—give a huge thanks to the rest of the Technology Committee in particular echo all of John's comments from Chairman Vacca and his staff. You know, your continued commitment for oversight hearings around the Open Data Law, and your energetic support has made New York City the open global leader, and we hope that you can find someone to toss the baton to. Hopefully, they're here in the room because we really



do want continued leadership in the same way that you have provided, and we don't want you to disappear in any way, shape or form. Okay, good. Can we have that—that's stated in the record, right? Okay.

Good. [background comments] [laughs] We also want to give a huge thanks to the City's Open Data Team for their continuing to slog through the notable losses in staff that they've had. The city hasn't had a Chief Analytics Officer for seemingly half a year, and then we've also lost a senior member of the Open Data Team. So, that has—actually, I think it's reflected in this report in the way that there are some things that were mentioned in 2016 that aren't carried over into this particular report. I want to point out the—the great observations that were made in the 2016 report. There are three observations that I think if those stories were told in the same way that John articulated, that they would strengthen the—the report itself on an annual basis. It's important to have a good understanding of what—what's going on within the accounting that we don't have to go into the spreadsheets for. It's actually quite frustrating that this report didn't have some of the high level overviews that last year's report was so

intricate. (sic) We've had some preliminary conversations with MODA around why those things were left out, and hopefully this will be a good opportunity to take that feedback, and to make the 2018 report even stronger. It's--there's some weird and some oddities that we discovered with both DEP and FDNY. Both agencies are going to archive datasets, and--and nowhere--well, FDNY added a column in their spreadsheet that say, hey we're going to archive this, but DEP didn't indicate that they were going to be archiving those data sets. So, in the 2017 report-- Well, let's see. We're in 20--the 2016 report indicated what datasets would be archived. This current report doesn't indicate that, and I would really want there to be greater clarity around that type of archiving because it impacts directly our data usage. You know, we don't want to all of a sudden see that ta-da, the data set that you are using today has now disappeared and now we have to go find another one. There were--the comments around FDNY that was--I--I can--I want to echo even further. It's such an odd thing to have an agency support-- submit and support a spreadsheet that says that this is our audit with such a sheet--a remarkable sheet

missing, and—and this was about legally mandated reports. So, how can an agency miss that a listing of legally mandated reports in an audit. That is just shocking, and I don'—I don't know how MODA would then go back and get FDNY to—to approve that report or clear up that report. You know, it would be shocking if the FDNY doesn't have to do any legally mandated reports. I just can't imagine that, and echoing all of John's thing, you know, statements around what is an improved examination and verification process?

How have the Open Data Coordinators improved their work, and then lastly a narrative around how are the different agencies moving open data or moving their data practices and open data from insight to action?

And the last call is for MODA to make it clear for these reports to exist on the Open Data Portal as well on their website. In regards to

recommendations, we worked with the Parks Department in in 2016 to do a user centered data release

process, which is three phases, and we've testified about this process before to user-centered or human-centered data release program, which is able to do three things: It's able to do a research and

discovery phase to find out exactly who were the data

users. The second phase is to do some user testing, and then third is to do some initial deployment. So, that way we can do some iterative work. This type of user-centered data release required a number of engagements, and as you heard from MODA, they would like more resources to be able to do that type of engagement. We call on the Council to figure out how MODA can get those types of resources to do this type of engagement because when we do the engagement before the data is final on the data portal, it saves us the end-user a lot of time and energy of having to then munge the data and clean the data, and go through the-kind of like a-a recursive process to fix the data, and to fix issues with the data, and there's nothing more frustrating than every month to send a letter saying hey, can you fix this data knowing that the whole process is automated and that means that there's another batch process that has to fix the data, and it's really frustrating. So, not to belabor you with the technical issues that we've experienced, we've been able to verify that the-MODA's recommendations on better citywide compliance. The-what is this six bullet points is completely spot on. When-this on page 3, so the-the issue around

every agency is unique. Executive buy-in is-is crucial. Open data is an opportunity to teach users how city operations work. Open data drives data governance analysis. Open data—people who use open data must be well networked, and I would also like to add trust, and lastly, organizational knowledge is a challenge or retaining that organizational knowledge is-is a challenge. We've been fortunate to work with the Manhattan Borough President and the Manhattan Community Boards to navigate this process, and to build those types and to essentially verify this recommendation. We're at the point where we have four classes that we can offer Manhattan Community Boards. We need the resources to get out of Manhattan so that way we can do trainings in the rest of the four boroughs. We're ready to do so. We have a very solid understanding of how to leverage the—the insights that MODA has been able to generate and apply that directly to community boards, and, you know, we're—we're—we're ready. We need—we need financial support to do so. We understand that we're entering in some odd fiscal time period but we're—we're ready, and we're willing to offer that to MODA and to do this partnership—to do this as a

partnership with MODA, but we need the financial resources to be able to hit the other four boroughs. We've been able to scrape by for Manhattan, and that's a temporary solution, and that funding runs out in December 31 of 2018, and then we'll be back to the original scenario where we have no money to essentially empower and train community boards around open data.

CHAIRPERSON VACCA: Are you a—I'm sure I know the answer, but are you a non-profit?

NOEL HIDALGO: Yes, we're—we're fiscal-

CHAIRPERSON VACCA: [interposing] A 501(C)(3) or that?

NOEL HIDALGO: Yeah, through the Fund for the City of New York.

CHAIRPERSON VACCA: Okay. Well, we have MODA still here. We have DOITT here. This was exactly what I brought up, and we have an organization ready to go, and we have boroughs that are under-served. So, I'd like you to take it back, and I'd like to have an announcement soon. I don't think this is a big deal. I'm—I'm—we're not asking for Fort Knox.

NOEL HIDALGO: No.

CHAIRPERSON VACCA: We're asking for an allocation to—to train community grassroots people on how to use open data to enhance their ability to monitor city services in their districts. So, I'd like—who from MODA can take this back? Who would be the—who would be someone here today who could volunteer to get back to me, and let me know if this request is something that's doable?

STACY GARDNER: [off mic] I'll do that.

CHAIRPERSON VACCA: Would you Stacy?  
Thank you. Stacy Gardner will get back to me.

NOEL HIDALGO: Great. Thank you. I want to also highlight one of the points in the spreadsheet, which John echoed, which apparently Craig has told me before this hearing has a very positive outcome, but there's a comment from the New York City Department--Health Department, which is someone from the Health Department leaving a comment on the Open Data Portal for FDNY to say: We are interested in using the EMS dispatch data that is provided on the NYC open data. However, we would really like more detail to the data if possible. Who should we contact to discuss this question? Thank you! How that story is resolved I think is critical

to answering the value of open data. In the response that FDNY provided, it said that their request was denied, and so once I mentioned to Craig, it was kind of crazy to hear that there's actually a positive story that comes out from that denial. But, you know, this is speaking to the value of the Open Data Program. It's speaking to minimizing FOIL requests. It's speaking to interagency collaboration, which is something that we've been asking for and clamoring for around open data for the last four years. SO, I wish that that that story, if it had a successful outcome was--was testified in--in the report and we hope that MODA will get the resources to be able to--to essentially lift up those types of opportunities across the different agencies, and we think that this is--these types of reporting is--is kind of like the baby steps for the types of resources for--for MODA and the Open Data Team to--to help be catalysts for other agencies to use each other's data, and to make our government services more efficient, and so with that, this is kind of a rough translation of what I've given to you in writing. We're always looking to combine MODA's objectives into our work and, you know, we are very thankful for all of their efforts



to essentially lend their brand to our work, and—and figure out ways for collaboration, but I have to echo this not a permanent solution. We have very temporal funds right now for Beta NYC's Operations, and we really need the support to move it forward, and MODA definitely does need the financial support to move forward. If they can't mention that, I'm going to stand here and say that they need money for it. So, thank you.

CHAIRPERSON VACCA: Oh, I—I-okay.

NOEL HIDALGO: I—just to echo that one last point. DOB, though John said that DOB did a great job, DOB mentioned five times in the—in the report that said: We do not have the resources to publish this on the Open Data Portal. So, that means that there are five other data sets that could be released by DOB that don't have—that they don't have the financial resources to release that data.

CHAIRPERSON VACCA: I thank you. A couple of points. When you mentioned the Department of Buildings, I—and I'm—I've been a person always—I've been a person frequently at odds with the Department of Buildings, but I do have to say that they really set a standard for transparency years ago

with the BIS system. Building information System, easy to navigate. It's very transparent as to what permits are filed, permits denied, why. So, they kind of set the standard with—with transparency. Now, they need more resource—more resources for open data, and I think you—you pointing that out is very important. I think we've come to the conclusion that the resources are not there. So, that was important to know. Thank you.

NOEL HIDALGO: Thanks.

CHAIRPERSON VACCA: Anything else?

NOEL HIDALGO: Well, actually, as we were discussing this, New York—I want to give a shout out the Planning, New York City Planning. They just released an updated version of ZOLA, which is completely using open data and using open source, and I look forward to heading back to my office to play around with that, and that's I think their third tool, and that's a team of three people that's essentially making open source technology in New York City government and doing ways—and moving very, very quickly.

CHAIRPERSON VACCA: And you are aware I think that tomorrow at 10:30 the Technology Committee

will be meeting again because we're voting on the Algorithms Bill that I proposed, and talking about data. Data is what makes algorithms, and algorithms, determine agency policy. We're trying to get greater transparency there, too, and open data is an important part of that. So, thank you for all your leadership here. We appreciate your advocacy all along and you provided it--both of you have provided our committee with needed--needed technical assistance. So, thank you. Any questions, Council Member Adams?

COUNCIL MEMBER ADAMS: [off mic] No.

CHAIRPERSON VACCA: No more questions.

Do we have further speakers? No. There being no other questions from the committee and there being no further speakers, it is now 2:10 p.m. and this committee hearing, our last hearing of the Technology Committee for 2017 is adjourned. [gavel]

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 December 20, 2017