****

This text is being provided in a rough draft format.   Communication Access Real-time Translation (CART) is provided in order to facilitate communication accessibility and may not be a totally verbatim record of the proceedings.

****

4-1-15

Captions will be displayed on this screen.

LISA MATTHEWS: Welcome.   Welcome to the third Department of Transportation open data forum.   We welcome everybody here.   We're glad you're here.   We have a full house, a lot of people coming from the public as well as from the city.   And we have some wonderful speakers who will tell us all about Mind My Business, and then Beta.   We look forward to hearing from both of them.   But let me first start by introducing our Chief Technology Officer Cordell Schachter.   (Applause).

CORDELL SCHACHTER: Thank you, Lisa.   By the way, this is Lisa Matthews, our Director of public data.   We take very seriously the city's open data law, and intend to comply not only with the letter of the law but the spirit of the law.   For us to do that, we are trying to establish and maintain a dialogue for the primary consumers, or the initial consumers of the city's open data.   And a lot of our activities over the next couple of years are going to be dependent on what we hear, and the priorities that we set and

the ways that we do things, we're going to try to base on the feedback.

We start -- just a quick story about two or three years ago when we first published a map showing the city's parking regulations, it was required by the city council.   They wanted an interactive map.   So we published the map.   And about two weeks after it was first made available, I got a call from a developer who said I want to start a small business showing parking around restaurants.   I want to sell that service on a website.   It was really -- he wanted to sell that service to restaurant owners all over the city. He picked out a couple of restaurants he thought would be interested.   He was sitting the Hagstroms.   At that time he had a text finder.   He said he'd been at it about 8 or 9 hours crossing off one by one on his Hagstroms map. He said this is crazy; they have to have a data set.   What is it or how is it you can possibly give me access to data set.   I said okay.   He started to argue. I said okay.   And he stopped me for a second and said did you really say okay?   I said yes.   I can't promise when adequate.   We're making the data available anyway.   You seem like a primary consumer and about five months later we posted the files of basically two files you need to match up to unlock all that data without having to -- from there up until four o'clock today -- so things change and things stay the same -- I got a call from the developer in Los Angeles who said that he's had no cooperation dealing with Los Angeles, and he's heard that New York City and New York City DOT in particular are friendly to developers who want to use the government's data. And I had pretty much the same phone call that I had two or three years again because he's just not that aware.   I told him coincidentally we're having an open data forum, and he thought about flying himself out here.

Just graduated with a master's degree and is interested in doing what probably a lot of you are doing in terms of creating applications using government data.   All over the country people are interested in the people's data.   And we want to find the best and most effective and friendliest ways to liberate that data to make it available to the people who pay for it.   So I thank you all for your cooperation, for your partnership, and I look forward to learning together.   (Applause).

LISA MATTHEWS: Just so you know that's our quick agenda. We're going to go through the welcome, introduce the city representatives. We're going to have open data at work, have our public speakers, and then public interest transportation topics of interest, questions from the floor.   At that point we're excited to hear all of your questions, inquiries, comments.

Let me continue on.   Let's start at the far end of the table because we have Albert Webber here.   He is civic engagement manager from the Department of Transportation -- transportation -- Department of Information Technology and Telecommunications and oversees the office of Applications and Technology.   (Applause).

ALBERT WEBBER: Hi.   My name is Albert Webber.   My title is Civic Engagement Manager but primarily I manage the open data portal, to a lesser extent development portal.   Over 1300 city data sets from a variety of agencies and developer portal allows you to connect to a number of APIs directly from the city.   Also part of my role there is making sure agencies are in compliance with our open data law which is local law 11 in 2012.   It's a lot of traveling and talking to agencies working with them to find out the data they own, what can be released and just making sure they're overall

compliant with the local law.   From a civic engagement perspective, taking feedback, going to events like this where we can kind of find out what the data the public is looking for, and how we can take the data already public and make it more usable and user friendly.   That pretty much sums up my role in the office of applications and technology.

LISA MATTHEWS: Thank you, Albert.   I'm going to come over here because I think you can see better.   Is this working?

Now we're going to -- Hassan, would you like to say a few words and introduce yourself?

HASSAN ADEKOYA: Thanks for coming today.   It's great to be able to be in a room full of civic technologists and receive information from you as well as we give information to you.

I'm responsible for process of GIS data for public dissemination as well as DOT.   We look forward to a lot of feedback from you.   For the best, and sometimes we actually issue corrections as well.   I think it's been a very fruitful partnership so far.   We look forward to continuing.   Thank you.   (Applause).

LISA MATTHEWS: Now I'd like to introduce you to Aileen Gemma Smith.   She will be speaking about how she uses city data, and will be speaking on Mind My Business.   (Applause).

Actually, I'm going to add in I just want to let you know we have added a closed caption panel here for anybody who would like to view it. We're very pleased to have offered that at this open data forum.   One thing to keep in mind, when you do ask questions or when you're speaking, please make sure we have one person speaking at a time.   Speak loudly and

clearly.

AILEEN GEMMA SMITH: Lisa, thank you. First off, thank you for the opportunity to share tonight, I sincerely appreciate it. I'm Aileen Gemma Smith. I'm the founder and CEO of Vizalytics Technology. I'm here tonight to talk about Mind My Business. But before I tell you about that I want to give you a little context of why we do what we do.

This is what my street looked like the day after Hurricane Sandy. It kind of had an impact on me. You're all probably wondering what does this have to do with open data? This is a local business that we've known for over ten years. It's family owned business, she met her husband there, the little girl is growing up there. The storm hit her really hard. That hard. What this illustrated to me was how significant of a digital divide there is because open data can represent a huge opportunity. But not if you can't get to it, or you don't know what's there. So what we saw is the potential to say there are folks that need information. But it's not there in a way that makes sense for them. So I said I want to do something to change that. Because a City of New York is an amazing place, it's a city of opportunity. And it's a city that's got over 1200 different data sets. So we came up with something called Mind My Business. What we do is let shop keepers know everything that's happening outside and around them that can make a difference to the bottom line. The very first thing that we did before we looked at the data was I went door to door and I asked shop keepers what's it like, what's frustrating? I wanted to understand where they were coming from. This is the feedback that we had from them. I trust you. You understand I'm busy. Just tell me what I need to know. That's part of the

opportunity in open data.

These are the problems I heard from local shop keepers. We're frustrated. We don't understand what regulations are that may apply to us. We don't understand when we may be losing foot traffic and we don't understand when our business might be interrupted because for example the street needs to be paved. Could you give me a heads up to that because that could make a difference to us. That's how we came up with Mind My Business. We use over 2,000 different data sets just at the city level and I want to show you how simple it is for the local businesses.

Sorry. This goes beyond New York City. Thanks for the shout out for Los Angeles. We're also working on this. We have close beta in Chicago, San Francisco and Seattle. This was a local shop that was an alpha tester in San Francisco.

When a business signs up with Mind My Business they just need to put down the business name and their E-mail. And their address. That's it. Then they start to see this. These are business specific notifications. They're everything from here's where you find, to here is what's upcoming with construction to here are problems reported in nearby shops. Folks love this. Think about who this data can make a difference for. If I'm a local shop keeper I'm not a GIS expert, I'm not a Python developer. I'm not saying please let me go fetch this. This is what I care about. I have a line out the door and I want to know if there's going to be a problem next week. That's what drives these alerts. We also made it super simple for businesses to get details.

Risks on roads reported nearby. You can share this with your staff.

Shop keepers love this.   It's simple, specific, telling them information they need to know instead of having to mine through lots of difficult information or not knowing where to find it, and then it kind of gets out of their way.

This is another notification.   We have a map view as well because some folks like to think about it in that way.   But for the most part the feedback that we've had from the over 700 businesses throughout all five boroughs has been I like this because this is simple for me.   This makes sense in my context.   You've actually honored me.   And you've shown me this is what I need to know that's specific to my business.   And that's part of the amazing potential of open data.

We've got all these different data sets, amazing folks at the various agencies here in New York City, at DOT, at 311, et cetera that are here and want to work with us and want to partner with us.   What else can we put out tomorrow?   How can we improve them?   Their hands are open for listening and for understanding what we can do, and also for calling out like wait, there's an error, or this didn't update.   So definitely, building relationships with that, but there's a huge opportunity as well.   For those of you that are thinking about how else can we use open data, or what are those other business cases, we love to put our hands up for thinking about what's that digital divide.   Our core user is the local shop keeper that's probably been around for 20 years.   He's the kind of fellow or she's the kind of woman who will say my entire life is inside these four walls.   Open data doesn't mean anything to me.   It's a jargon term.   But telling me there's going to be construction in two weeks is meaningful.   Letting me know a sidewalk is coming down, it's meaningful.   Let me know there's a water

main break nearby, that's meaningful.   It's simple.   That's where we see the potential.   Thank you.   (Applause).

LISA MATTHEWS: How exciting is that?   I think it's wonderful.

We're learning.   So our next speaker is going to be Noel, he's going to speak about Beta Talk.   And let me just say now we've had our chance to see open data at work, and now we're going to have a bit of conversation on the discussion groups that exist.   (Applause).

NOEL HIDALGO: Thanks.   First I have to say thank you Department of Transportation.   This is great.   We wish and we hope to see more agencies do these types of open data hearings or at least town halls.

We're just going to balance like all things in New York City.   It's okay.   I got it.

I just want to be brief because I know that the community introductions only have a few minutes.   For those of you that don't know I'm Noel Hidalgo.   We're the city's open data and civic technology community.   We host weekly events and advocate for open data and civic technology to improve everybody's lives.   We create a safe space for government and the public to collaborate.   This being the third forum, this is a great example of what we're doing.   We advertised today to our community members and I hope -- I see a lot of faces out there.

The problem that we see with open data in general is that most open data conversational spaces are formal, just like this is.   I know this is more informal than what we're used to, than presenting hearings to city council. But this is still a very formal space.   And when we're answering questions

around open data we routinely find we think those data responses should be open and available to everybody else.   And we also feel that we've also seen the community can help answer each other's questions, and remove some of the burden placed on government to engage in those questions.   How many of you have used the comments section in the city's open data portal?   I just want to see.   Great.   That's six hands.   How many of you didn't know that that existed?   A lot more.   Right?   What we want to do is remove that burden, not kind of.   We want to remove that burden.   We see that Department of Transportation's data is really at the core of the city.   It is the streets and sidewalks.   It is the place that we go when we leave our front door or the last step that we go into our office.   And there are a number of nuances with the data that needs to be documented and socialized.   Cordell was just giving the great example of talking about how you need to combine two data sets to get the street with the parking regulations databases.   And the city's current data platform, no offense, Albert, but it just doesn't provide for peer to peer problem solving.   That's a big problem we found.   How do we increase usability?

What we'd like to do is to have an online collaborative space.   We've kind of created that through this platform, open source platform called --

LISA MATTHEWS: We're all thinking about that.

CORDELL SCHACHTER: Remember we're self-insured.

(Laughter.)

NOEL HIDALGO: The city's event liability insurance, right?   What we've created through BetaNYC, through the hands of several people in this room, is this forum called talk.beta.NYC.   It is that online public space for

engaging and solving kind of the -- it's a place for you to ask questions and have them answered by the community, and hopefully through government or with the assistance of government.   So we put together talk.beta.NYC that organizes just a few little pieces.   So we have a general discussion forum.   We have a developers' corner where if you're interested in asking a more technical focused questions you can go there.   There's a jobs board. And then we started to develop this thing called working groups.   And at this time we have -- there's a small group of people that are looking to improve the open data portal.   Last year we got the city record, a piece of legislation to get the city record online.   And the city record is the city's newspaper that publishes all events, and so we're now working to put that into a format.   This is a place where that community organizes and discusses.   MTA service alerts.   There's apparently the MTA has a really detailed service alert feed that's out there.   But it's not in a machine readable format.   So we have a team that's working to kind of disembowel that data feed so that way you can get excellent push notifications.   Recently the city put out a new kind of search website called DORS.   We've stood up an online location for people to kind of play and participate and engage in -- DORS is an open source platform.

What we're proposing is that we take kind of the enthusiasm and activity that happens in this room and kind of put it online with the participation of the Department of Transportation as well as the participation of the city's open data and civic community that we come together and we use -- instead of using a Google group to ask questions about the city's open data that we use talk.   There are a number of

organizations that already have precedent of using peer groups to solve these problems.   The NYC big apps, the MTA, the Connecticut transit, and who is the fourth one that I'm missing?   I'll remember as soon as I sit down. But essentially they're using Google groups to engage -- oh, did I say Citi Bike?   MTA, Connecticut transit and NYC big apps are all using online groups to kind of talk about their data to engage with the community.   The last time we had one of these things they said, Department of Transportation said, Noel you have to get up and talk about this, and see whether or not there's consensus to use such a portal, or use such a forum.   So with that, I encourage you to use Talk as a venue to engage with the city's civic hackers broadly, to engage with the city's civic technology community broadly and encourage the Department of Transportation to work with us to figure out how to harness talk as a way to answer the open data questions that this community has and find ways to continue this dialogue that we have that we've successfully had in a way that's online and continuous.   With that, thank you.   (Applause).

LISA MATTHEWS: Thank you, Noel.   Yes, we are considering various ways how we want to move forward in creating discussion groups. We're not quite there yet.   We plan to be but this was a forum to have more ideas posted.

With this we will continue on.   We're going to open up the floor for -- there we go.   Now come on.   For the past couple of weeks I've been receiving inquiries from everybody.   This is the time we'd love to hear your questions, inquiries, and we'll have Al, Cordell, Hassan or I welcome your questions and we'll come up with an answer.   Just so you know, Cecilia

Goldman, is here. She's in my office. She's going to help us with the microphone. There are three. We want you to speak loudly and clearly and she will bring the microphone to you. Thank you. Who would like to begin? We got to have some questions here.

FROM THE AUDIENCE: Hi. My name is Deger and I'm building a parking app by using the city's parking regulation beta sheets and I was wondering do you plan on improving the API? Right now we have to do the files, and I ran into some troubles with it. I was wondering do you plan to improve it in the near future?

HASSAN ADEKOYA: I will ask if you can elaborate on improving it. However, the idea is also a shapefile that you can also download. I'm not sure if that's useful for you.

FROM THE AUDIENCE: Yes. I saw the shapefile and I opened it up. I saw different formats for the state system and I used software to convert the system into latitude and longitude. And the parking signs that I found are not matching to the real parking signs that I see on the streets now. There's a difference between the data set you have and the parking signs that are on the streets. That discrepancy caused some problems with my software.

HASSAN ADEKOYA: For the discrepancy that you have, I would ask that you send those to parking regulation feedback at DOT.NYC.com. It will be investigated and, if needed, verified in the field and corrected so in next month's update it's corrected. We try as much as possible to have confidence in the data but it needs verification. So if you notice anything that is wrong -- just to give you insight into the regulation data set, it's an

ETL of 29-year-old system.   It's been modernized in a couple of months.
So in a year or so we should be able to see a modern looking management
system.   But for now we're relying on this ETL as converting based on
distance from curb -- it's an algorithm that we feel confident about but if you
notice anything that's wrong, it may be a problem with the ETL.   So the
way we can improve the data set is by you reporting it so that way we can
correct it.   Take a screen shot or report exactly the intersection that you
have issues around and we can send it for fixing.

FROM THE AUDIENCE: Thank you.

HASSAN ADEKOYA: Parking regulation data feedback.

FROM THE AUDIENCE: Hi.   My name is Henry.   I'm part of the
team that Noel mentioned that is working with the MTA service alerts.
There are some inconsistencies in the way the data is formatted.
Particularly that some things are surrounded with tags and other things
aren't.   I was wondering if anybody knew who I might be able to reach out
to for perhaps changing a little bit of the consistency to make it more easily
parsed.

HASSAN ADEKOYA: What data set?

FROM THE AUDIENCE: It's on the -- called MTA service alerts, at
the very top MTA all agency service status that refreshes every minute.

>> The MTA service alerts.   XML end point.

FROM THE AUDIENCE: It's an XML document that gets over
written once a minute with the current status.

HASSAN ADEKOYA: That's on MTA's website.   So MTA has a
very rich developer community on Google forum.   Hopefully in the

future -- developer community is very responsive. I would definitely encourage you to look into their Google forum. Somebody from MTA routinely checks that website. If service is down they will let you know as well. So I advise you to post your comments there, and you are bound to get a response.

FROM THE AUDIENCE: Thank you.

ALBERT WEBBER: I know the MTA feedback is a question I'm sure DOT receives all the time. The MTA is technically not a part of New York City, New York City government. So when that feedback comes in, that is separate from us and New York City government.

FROM THE AUDIENCE: Hello. My name is Richard. I was wondering if there was any way when the Department of Transportation installs a new sign conducting a traffic flow, is there any way to find out the exact date the sign was installed and why it was installed?

LISA MATTHEWS: I'm sorry; could you say when -- Department of Transportation installs a new sign?

FROM THE AUDIENCE: Conducting a traffic flow.

LISA MATTHEWS: Okay, the Traffic regulation sign that has been changed.

FROM THE AUDIENCE: Any way to find out the exact date it was installed and why it was installed?

CORDELL SCHACHTER: The date is probably there. Whether it's accessible now or we're servicing it now I'm not sure. I don't believe the sign asset database has the reasons why the signs were installed. For signs as well as traffic signals, there are engineering studies done before any of

that is installed or existing conditions are changed.   I don't know that we publish that as a matter of course.   The information may be available through FOIL.   So if it's something that's important to you that may be the avenue to try.

FROM THE AUDIENCE: Hi, my name is Lindsay.   I was just wondering if anyone could just describe the history of trying to engage the design community around your initiatives.   Because data is only as usable as the screens on which they're displayed.   I'm a designer myself so I'm curious what the efforts have been like and the success.

LISA MATTHEWS: I was going to say to be a little clearer on the design.   Do you mean on the design of the street or the intersection?

FROM THE AUDIENCE: No, on these apps.   I'm a user experience designer myself so I'm curious what the success has been because developers are not the only people that make up developing a software product.

CORDELL SCHACHTER: As we mentioned earlier, this is our third public forum.   So if you establish either a conversation with us, other people want to contribute to that either in public or you E-mail us, we can get things going that way.   Or as Noel and Lisa have suggested, the various collaboration portal is available on the web.   Does that answer your question?

FROM THE AUDIENCE: Not really.   Traditionally open source communities had a hard time engaging the designers because we're not doing the technology implementation but we're helping with the adoption of the product or the service.

ALBERT WEBBER: I guess from the overall open data portal perspective, when we initially kicked off with our new vendor that we're using right now in 2011, we didn't do too much engagement for that actual design work.   Recently we've done a little bit more when it comes to visibility things.   One piece of feedback we've received, when you have that Navy blue background, having certain color fonts and things like that to help it be a little more accessible, we've tried to do a better job of doing that. But that is something that we would like to -- I guess you can look at it as the open data portal has set up a lot for developers at this point and we'd like to open it more.   So when you have people with design experience when it comes to data standards and things like that helping us with descriptions for data and stuff like that, that is something we plan to do in the future.   I would definitely say look out for things like that this summer.

HASSAN ADEKOYA: To also follow up, at DOT we routinely work with designers on staff.   But we tend to contract those out.   What we found in feedback that we've gotten is that maybe before the launch of some systems we should talk with some civic developers or come to some of those meet ups and have a discussion about the systems.   But we've gotten feedback in the past and it's definitely something we have in our tool to use when we release new systems.   This feedback is very important because we -- some of us are developers.   When it comes to developing applications for the public, it's a different insight.   We're beginning to realize that a lot with respect to the feedback we've been getting from the community.   We need the assistance.   We'll keep an eye out for that.

FROM THE AUDIENCE: Hi.   I work with the NYPD's crash data.

One of the biggest problems that I face when dealing with it is ambiguities and the location of crashes.   About ten to 15 percent of crashes that are listed we don't actually have a latitude and longitude for them.   We'll have a cross street, street one and street two.   But then it can be ambiguous as to what the actual intersection is because these streets might cross twice or might not intersect at all.   My question is what are the steps being taken to address these ambiguities in the crash data?

ALBERT WEBBER: I wish I had somebody from the NYPD crash team here.   From my understanding, when the crashes come in, for the most part in the open data portal you'll see the intersection.   So I think what happens if you have a crash midblock it will be assigned to the nearest intersection.   I guess there are some circumstances where that cross street isn't taking in the best -- that's feedback I can pass along to the NYPD crash analysis unit.   I guess I can try to relay that information back to them.   A discussion more like Noel's or that's feedback I can throw back up there.

FROM THE AUDIENCE: I have a related question.   I also work with crash data and transportation alternatives.   I work with Vision Zero, the DOT program.   And there are discrepancies between the NYPD open data portal and Vision Zero views.   I was wondering if there are plans to collaborate with NYPD and sort of what the plans are for updating that platform in the future.

HASSAN ADEKOYA: I guess my question is could you elaborate on the discrepancy?   We use the NYPD -- that's going to be a common routine line throughout this discussion.   But we use the NYPD data source for Vision Zero view.   Discrepancies that we are aware of are with respect

to fatality.   We're using it to consult the NYPD database to record those fatality numbers so if that matches what has been reported, that's the official record of fatalities around crashes.   Would you mind elaborating on discrepancies?

FROM THE AUDIENCE: Vision Zero data is under reporting fatalities.   But you're using open data portals?

HASSAN ADEKOYA: We're using the data for the crash data sets, just the crash that caused fatalities.   Just injuries we use the open data. Same feed.   The numbers should match.   Fatality, when it doesn't match at all, that's why we're using a database that is not -- that is the traffic system because it's just under reported.   We know that.   We're working with NYPD on that and they're working on improving the system, some of the action items that are reporting the NYPD to get better.   It's something they know of and they're working on it.

FROM THE AUDIENCE: I'm very glad I'm here.   Thank you for this opportunity to explore interaction which has occurred in the last 10 or 15 years which did not exist 25 years ago when I was working for the department of finance.   I retired ten years ago.   But 25 years ago I started an organization called Gismo which some of you may have heard of, it began to bring people from different agencies together so we could do exactly this kind of stuff.   Because too many agencies, we're isolated from people.   The tech people are isolated from one another and you couldn't do this but the data was there.   I have two questions for you.

One relates to the Mind My Business application.   I live in flushing. Oh, it's a real beehive of small businesses.   Most of the people running

those small businesses, English is a second language or not even a language. So many people speak only Korean or Chinese.   Is there anything being done to translate what you've got on that website so the people who speak only Korean or Chinese or Spanish, whatever language, can access it?   I have a second question, and then I'll shut up.

The second question is one of the things that have not happened in metadata, one of the things that I hope would have happened 25 years ago is that the city would see beyond its political boundaries.   Right now data is collected politically.   We collect it in the City of New York.   Cross from Queens into Nassau or from Bronx into Westchester or from New York into New Jersey and the data doesn't match.   Is there anything being done to get the metadata to work with similar organizations just outside of our city so that there are applications that go beyond the city limits?

AILEEN GEMMA SMITH: Thanks.   First off, I'll say this.   (Not English) I lived overseas in Tokyo for 12 years so I absolutely empathize how challenging things can be when they're not in your native language. We take that to heart.   We are working on simplified Chinese and Spanish. Because the content is dynamic, it's not just hey translate these phrases and make them work.   We do take it to heart.   In order to reach all, that means not only English.   We certainly take that to heart because you are leaving out huge swaths of the population.

In terms of metadata if I could speak very briefly from our point of view.   As we work with other cities we see some of those problems as well. I think folks working within the enterprise to do that are looking at ways to sort of normalize and see best practices across cities, like New York, San

Francisco, here is what we're doing that's worked really well but we need consistency. I defer to the team.

HASSAN ADEKOYA: Thank you, Jack, for coming. Before there was open data there was Gismo. Gismo allowed for city employees and private entities --

Again, it allowed for that collaboration and open discussion way before there was open data law. Jack supported this for a long time. 25 plus years before you retired. Thank you again for all that. We won't be -- all this discussion about data wouldn't be happening if not for this group.

With respect to metadata, it's a nonissue. There are no standards for it. There are no standards that I'm aware of for data that is contributed to the open data portal. I don't know if Al wants to comment on that. From our perspective we try as much as possible to describe the data as effectively as possible so that others can use it. Do we use fields or formats, federal standards or state standards?

ALBERT WEBBER: I guess I'll try to speak from the perspective of all agencies I've worked with in terms of collecting data. It is known that the metadata that goes out in the open data portal is not I guess the standard that people are generally looking for. In terms of that, in terms of those boundaries from the city to maybe Nassau County, or maybe New Jersey, not too many of those conversations have happened. I know there have been conversations with 311 here and in other cities; but not necessarily ones that border here and New York City. So there have been conversations with New York City and Chicago and San Francisco and

Baltimore about sort of standardizing those 311 service requests and things like that, same thing with restaurant inspections.   But in terms of that metadata overlapping local boundaries, there obviously have not been too many of those discussions.

FROM THE AUDIENCE: I just want to pick up on the point that Lucio and the woman from MTA mentioned about the collaboration, the connection between DOT data and NYPD data.   I was thinking it could be great to hold a forum with both agencies so that both members of the public and the agencies can see what the process is for you to collect data and just how the data flows from one agency to the other.   If that's something you'd be interested in, I think you could get a lot of interest from the community.

FROM THE AUDIENCE: I have a question and a comment.   First of all, one comment, one best practice we're trying to promote in BetaNYC, whenever we create a map we encourage people to have a link to the data. It would be great -- we encoded that in the community portal for BetaNYC. It would be an app that the DOT puts out that they get the link.   When they click on that they can see the responding data and what was done to that data, how it was drawn.   So at least we don't need to do the same thing over and over again.   Often times what we see in the communities, they would build scrapers, which brings down a DOS effect on your servers.   If you put the link to the data, people won't do that.   We highly encourage it.

HASSAN ADEKOYA: On the strategy, the Vision Zero view for the data, we should do what you want.   And it's updated.

CORDELL SCHACHTER: It's really due to the conversations that we started here.   We realize to supplement the city's open data site which is

up on the screen on the side of the room; we wanted to make our data available in context.   So while you're looking at the app, that's the thing you're looking at, there's no uncertainty either by going to our data page or the city's page that you're going to grab the right set.   Our standard will be going forward that as we create these apps or version them forward we're going to create links.   In the case of something like Vision Zero view where the data is refreshed about monthly, we'll include the date and the link so you know what you're getting.   In other data sets where the data literally refreshes minute by minute we'll include what the SLA is.   You'll be seeing that more and more from us going forward.

FROM THE AUDIENCE: Just a follow-up on the standards. People might be interested, there's a standard that is being currently promoted right now called buildingpermits.org.   We're a member of that standard group along with companies like Acela.   The great thing about that, we're trying to mimic what happened with GTFS which allows directions to get here.   Google and all these other direction finding things. We're trying to standardize building permits.   Which directory translates -- things like what Aileen was talking about.   Right now it's in a comment period and if you're interested in giving your comment and the directional standard, go to bldgpermits.org.

FROM THE AUDIENCE: Hi.   I'm Austin, a researcher with the social policy simulation center in the college of Staten Island.   We're trying to work on transport was a system installed by the Department of Transportation to monitor traffic flow in the city.   They also have another partner providing some data.   But is there any API currently available

getting that transport data out to the public to use for research purposes trying to simulate if another event such as Hurricane Sandy were to occur to try to figure out the best way to evacuate people out of the city would be or something along those lines?

CORDELL SCHACHTER: I don't know if there's somebody here from DOT who wants to answer that.   I'm not sure what's available directly. I know we have a number of projects in motion, and concepts with universities such as RPI and -- specifically about systems that would be maintained, whose primary purpose would be to help people during crisis.

FROM THE AUDIENCE: Right, but would that data be available in its raw format for research purposes?   If the college of Staten Island or CUNY wanted to go forward, is that something we could look into?

CORDELL SCHACHTER: We're the ones exposing it, then you'd use the same standard about making the data available -- available in the city's open data platform, but also in context with whatever the application is.

FROM THE AUDIENCE: But is that something -- is the DOT having a road map to release that transport data to the public at this time?

CORDELL SCHACHTER: That's the part I don't know.   I'll find out.

FROM THE AUDIENCE: Okay.   Looking to traffic studies vehicle counting's, that data is published in a pdf online on a server.   Previously to get that data you had to manually go onto the DOT's website and download each pdf one by one.   To alleviate that problem I had to install a scraper program.   Even though they are quite dated from 2007 and 2006, is there

any way to put it on the open data portal?   They are there but they're numerically numbered but you have to click on each segment within the GIS application to pull them up.   You're then linked to the pdf to download it. But if you have to download 50 or 100, you have to manually do each one.

CORDELL SCHACHTER: It's part of the road map for current and future accounts to be made available in raw data form.   We don't have funding or research right now to go back.   What we talked about at the previous two meetings is probably worth saying again today.   We commit to a couple of things.

One, we're going to obey the law and all the data that's supposed to be released by 2018 according to the law will be released by that deadline, informally ahead of the deadline.   The second thing is where the data and the ways of servicing it or making it available are consistent with our purposes, government purposes, because that's how we're funded right now, and then we'll also make a commitment to do that.   Where additional things are being requested by the public, by the open data community, I can't commit to that without the resource.   But one of the reasons why we're having forums like this is we want to hear if there's a demand.   If it's more than just one researcher or one developer who wants it, it's easier for me to go back and get resource for that even if DOT and New York City doesn't have a governmental purpose for it.   But it would be an important source for academia, for the business community.

FROM THE AUDIENCE: Thanks so much.

FROM THE AUDIENCE: Hi.   Really general question to follow up what he just said to what extent do you track how everyone is using your

data? Is there any function you're really excited about and it just sort of underperformed? Besides holding these forums, how do you know people are using your data?

CORDELL SCHACHTER: How do we know about people using our data? We get a certain amount of activity from 311, through E-mails to the agency, through contacts, Twitter, other social media as well as to these E-mail addresses and phone calls that we get from time to time. There's no -- we don't have a formal system of metrics. We're not recording it as a KPI, key performance indicator. If that's something we'd need to do in the future, something City Hall would require us to do, there's potential for that. But we're busy with what we have. And it's enough of a challenge for us thinking about our internal customers and primarily DOT's application development which right now we support about 80 custom developed applications for internal DOT customers that were created over the past 20 years. Some are pretty archaic technology. So we have a lot to do just keeping the lights on. As we refresh those we're trying to defer open source and project management as well as using our budget creatively to be thrifty and value based but also not to be slow as we have in the past. So we're looking at that. Looking and providing direct service to the public in that way, so Vision Zero view, I ride NYC, we've just been in that business about two or three years so we're still learning. It's one of the reasons why we want this dialogue going forward.

ALBERT WEBBER: From the perspective of the open data portal, what we're doing with the data you find the DOT data feed, we're linking directly into that data for consistency. Our analytics page is available so

you could see the number of views, number of downloads for a lot of these data sets.   I believe it's data.cityofnewyork.us/analytics You can find those charts and frequently searched terms.   You can see those views and you can follow those trends.   I think down to the LE level.

HASSAN ADEKOYA: We don't have the metrics directly.   Just for looking at E-mails that we get, Lisa and I get so many E-mails daily.   Each E-mail involves a fair degree of research.   The data set that we have is the parking regulation data sets.   There are questions about accuracy, also data usage.   We welcome those as well.   We are also coming up with more formal metrics for communicating requests for clarification to DOT.   That way you can track the systems as well.

FROM THE AUDIENCE: Hi.   My name is Tim.   I have a comment that I'd like to second the notion about having multiple city agencies get together and have a forum like this.   Because sometimes we don't even know who owns data.   Certain data sets.

Another comment is I think it would be great if all the GIS data was all published with the same projection information.   I think the NYPD data produces it in the New York, Long Island, x and y coordinates.   Not to get too technical.   Other groups do it in different formats.   You have to keep converting data to be able to combine data sets.   And finally, I was looking at the, I guess, the moving violation data.   I don't know if it's a DOT data set.   But is there in general a feedback loop when someone is doing data quality review on these data sets to provide feedback to the data provider to clean up the data or some -- even if it's historical information?   Because I found there's like tons and tons of errors in some of these data sets.

ALBERT WEBBER: The parking violations data is from the Department of Finance.   When that feedback comes in, what we have at all agencies, and you can see some of this in the technical standards manual in regards to open data are open data coordinators at every single city agency. We'll take that feedback and route it to those agencies.   But I think when that feed went out, it's updated on a monthly basis, the idea was there are errors in the data.   I don't know if its fat fingers and you have the ticketing officers entering in the inaccurate information.   But it was released the data exactly how it comes into the system.   I know if you look you'll see some parking violations dated in 2000 100.   I think that was a decision just to release the data exactly how it comes in.

FROM THE AUDIENCE: I guess one of the major issues -- someone had done -- there was a blog about parking issues, and I found that I wanted to remove -- I wanted to separate the data from personal automobiles and commercial automobiles.   I said whatever this analysis; I don't want to count trucks because that doesn't play.   And I found that most of the trucks were listed as passenger vehicle, even though it was a freight liner or a Mac, that's not a fat finger; that's just bad data.   That will never get changed, or it will just keep getting badder and badder.   My question was could you go back and fix that freight liner and change it from a commercial vehicle to a commercial vehicle?

ALBERT WEBBER: I'll route that back to the department of finance.   I'll pass that along to them.   I can only -- from the do it end --

FROM THE AUDIENCE: I don't know what they're doing now. I'm gone for ten years.   (Laughter.) And it's pretty hard to find out anything

in our electronic system. You try to find out something on the web from Department of Finance, you don't get anywhere. You don't get E-mail addresses for anybody.

ALBERT WEBBER: From the DOITT end, my role, we collaborate and bring this data all together in one place so I have an understanding of it. But again I'll have to forward that to the Department of Finance. I feel like something could be changed about how that data is classified.

FROM THE AUDIENCE: Hi, I blog about autonomous cars. Ipso facto, you guys have any awareness around any data or methodologies or models coming down the pike where you may be looking to use in the near future that can frame the impact of that technology on any new construction, any new projects in general you guys are undertaking?

CORDELL SCHACHTER: I don't believe there's anything that we're doing independently. DOT is once again a member of the intelligent transportation society of America whose chair is from the Michigan State Department of transportation. And they've created an autonomous vehicle test bed. I believe around Detroit, a consortium between government, GM, and some other private companies. I think the next annual meeting of the ITSA is in June, May 31 to June third in Pittsburgh, Pennsylvania. One of the main tracks is about autonomous vehicles. They'll publish papers. They lobby Congress right now. They're trying to keep the spectrum allocated for infrastructure to vehicle communication and vehicle to vehicle communication, the 5.9 megaHerz band out of public domain for general use so it's dedicated for vehicles and transportation information. We're looking to use it here in New York City not only for infrastructure vehicle

IV, and vehicle to vehicle, v to v, but also what's being called i to x.   In our case it would be pedestrians, to give better awareness of what's going on around them on the street.   Also to make information available for various groups of pedestrians who could use more help, like vision impaired, mobility impaired, to help know what's going on around them, and potentially to provide tools on smart phones and technology.   We definitely will be part of that.   But right now the investments are being made outside New York City.

FROM THE AUDIENCE: Thank you.

FROM THE AUDIENCE: Hello.   My name is Vahesh.   I work with some maps and bike lanes.   I was wondering if the DOT has any plans to update the bike lanes or do a more regular release of them.   The last one I'm aware of came out in mid2013.   Do you have plans to update it?

CORDELL SCHACHTER: Anybody here from the bike group in DOT?   I can tell you there's work going on all the time.   We'll find out what the next major release will be and we'll announce it.

FROM THE AUDIENCE: The bike map, the 2015 bike map is now being printed.   That should be out soon.   But I'm not --

FROM THE AUDIENCE: It is in print.

FROM THE AUDIENCE: There's a new one coming out.   Every year a new one comes out.   That will be the latest but I know what you're talking about is different.   I just wanted to say 2015 is coming.

CORDELL SCHACHTER: If it exists in print we can make it.   That would be an easy lift.

FROM THE AUDIENCE: Please do.

CORDELL SCHACHTER: Thank you.

FROM THE AUDIENCE: Just one comment Cordell about your comment about the systems.   One other thing we've been advocating for at NYC is this whole concept of using services architecture, API.   As systems are retired and new systems are put in place, the interface stays consistent.   And that helps us developers.   Hopefully then you'll be exposed to safety so developers can use it too.   Speaking about API's, one great resource that is a great demonstration of a legacy system being wrapped in API is geoclient.   I know there's been some talk about releasing geoclient.   Right now it can be downloaded from DCP's website.   What some people are looking for is if they can also release the geoclient API wrapper so people like us -- one thing we often find when we clean up data is we geocode the data sets.   Often times we use commercial services like Google, which are very slow, very expensive, and not as really authoritative as what geoclient gives.   Since it's out there already, I think it will be another open data first for New York City if the most amazing New York specific optimized geocoders release to the public.   Because that I think will go a long way towards really -- the community cleaning up the data.   And for start up like us, even Vizalytics if I can say so myself, sometimes when we look at the data sets, they use the xy plane, and we have converted to wgsa, we don't have to do that, and all the other things we look at, what's the community district, what's the police precinct?   It's all there in amazing detail.   It would be great if that's also released.   Thank you.

HASSAN ADEKOYA: I'm going to take the easy question.   At DOT we have applications toward that approach.   We use services

architecture for architecture development.   The development is on.   It's one that we will do that will take two years to build just because of master business practices.   As they're released, they will be having an API which makes it easier when the time comes to release.

CORDELL SCHACHTER: You heard me say we're trying to do that in two years.

HASSAN ADEKOYA: The applications.   Around the APIs themselves, the key concern is while the application is built to support the business, government's use, when it comes to supporting public consumption; we need to discuss how to be able to build an application when we don't know what the variable use of the public will be.   So internally we're having the discussion.   It takes time to find the framework. As we're presented with ideas we'll be discussing that as well, in this forum as well.   With respect to geoclient, we love it as well.   We use it for all validations.   It's great that city planning works together with the API. They didn't provide the code, the source code for the geoclient but with the way city government is created, it's just a matter of time.   You have the community, please make it work.   Request, complain.   Requests.   Let's call it requests.   It's just a matter of time.   I definitely echo that.   Just being able to have a consistent way to be able to get districts, get xy coordinated planes.   And all the recent conversion communities converting it back is very easy.

FROM THE AUDIENCE: My company has a platform for building transactional mobile apps, mostly used in the commercial world.   I'm very excited about some of the data sets where we might be able to show that off

and provide a service, even better provide a real service to someone. I notice there are a lot of databases of work orders that are being executed by city personnel. Is there -- any of you have any thoughts about a database that might be very interesting to work against transactional app where it can sort of read things in like work orders as an example and make that data available to people who would be acting on it? Doesn't have to be work orders but since I have all the experts here, what might be an interesting API to exercise our system against?

CORDELL SCHACHTER: One of the things that we're talking about is making the service request information to 311 publicly available, and for multiple purposes. One, so the public sees what we're doing. Secondly, we're burdened by repeat requests. Two people who don't know that they've asked for the exact same thing, we have to answer separate requests. For us to join them together sometimes people who explain it a different way, it takes us a while to figure out the request. Where if the first request was transparent and easily viewed, perhaps you could follow it instead of a second one, so people are both letting us know it's serious and making our work flow a little more streamlined. We're working with do ITT right now in the engineering of 311 for that and to make that data available. So groups like yourselves -- you could get potentially the request but not necessarily the internal system work order.

LISA MATTHEWS: Just so you know, I don't know if you've looked at the NYC open data site.

FROM THE AUDIENCE: Just a bit.

LISA MATTHEWS: On the side screen we have both sites, the DOT

site available where it says data feed and the second site where they have APIs is on the NYC open data site.   I'll let Al to speak to that.   That's his baby.   I know in that site there is 311 service request information.   In that you will find complaint topics, and you will find work order numbers, and you will find location information.   And there's a multiple ways to look at the data.   You can map the data and correct me if I'm wrong there is an API there for that, correct?   Let me let Al speak about that a little more and you can start there.

ALBERT WEBBER: In terms of the work orders, I'm just going off the top of my head.   I think -- I don't know if you all are aware of open data plan, it's pretty much a plan that maps out data sets that city agencies -- it was created in 2013 -- top of my head, I believe Department of Environmental Protection may have a work order management module that is coming due pretty soon.   When we get that data, you'll be able to go to the open data portal and connect via API to get hold of that data.   In terms of other work order management stuff, I know some work order things come into 311 but I'm not sure if they all do.   But some can be found at 311 service requests.   Nothing else -- to your question, nothing else comes to mind in terms of that work order data.

FROM THE AUDIENCE: Can I also speak to that?   If you're interested in that kind of data, I'm going to make a plug for one of the projects on BetaNYC right now is the city record online project where we're trying to parse the information put out in the city record by the department of contract and administrative services.   That includes a lot of things like procurement requests by the city and them seeking contractors.   The ideal

is we have machine readable format that can then be accessed through an API and hopefully that would allow various subcontractors and people independent of the city to be able to go through a portal that would then connect them directly to the contacts they need.

LISA MATTHEWS: We are coming to the end.   We want to take one more question, maybe one or two more questions?   This gentleman has been waiting quite a bit.   Speak loudly.

FROM THE AUDIENCE: I was wondering are there any plans of sensors to keep track of the public parking spaces?

CORDELL SCHACHTER: We tested it.   And the answer is no. It's very tough in New York City with streets being plowed to keep the sensors in good shape.   One of the challenges that we had is managing the on street parking inventory without marked spaces.   There are various technologies that we looked at, both video as well as in street sensors. We're going forward with basically as many as you can fit in along a curved space that's been allocated for parking.   It makes it difficult we know to publicly advertise where spaces are available.   There are some developers working on density and being able to base on looking at the parking transactions, coming up with predictive models that would give drivers a better indicator of where they're more likely to find spots.   It's something that we want to do.   We believe not only would it satisfy parking customers but less traffic.   The place you're most likely to find parking.   But we just haven't found a good sensor that works with the whole system yet.

LISA MATTHEWS: One more question.

FROM THE AUDIENCE: This should be quick.   I work for via, a

ride sharing here in the city.   One set of information I haven't been able to find online is restrictions with Manhattan and potentially other boroughs. I'm wondering if that is information that's publicly available.

CORDELL SCHACHTER: The best database for our traffic control information, signage, for signage is our street sign asset management system which is one of those 30-year-old systems that we're in the process of modernizing.   It should be done next fall.   From that it will have a very powerful search tool that you should be able to search on a particular type of regulation.   And then see the result.   And in addition to -- again, the primary customer is internal DOT being able to manage the signs and traffic control.   It will have a public facing portal or aspect to it that should let you do that.   But right now it's probably very difficult.

FROM THE AUDIENCE: Thanks.

LISA MATTHEWS: Okay.   I think this will conclude our forum.   I just want to say thank you very much for everybody for coming.   Cordell, Al, would you like to say anything?

CORDELL SCHACHTER: Same thanks and we look forward to continuing the dialogue.

LISA MATTHEWS: Please visit the websites.   You can find a lot of information there, and the NYC open data there's a lot of information. Thank you for coming.   Thank you for the speakers.   Thank you for making this meet-up.   We appreciate it and getting all the beta talk people here.   We appreciate everyone being here and we look forward to continuing the dialogue.   (Applause).

****

This text is being provided in a rough draft format.   Communication Access Realtime Translation (CART) is provided in order to facilitate communication accessibility and may not be a totally verbatim record of the proceedings.

****