CITY COUNCIL
CITY OF NEW YORK

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

TRANSCRIPT OF THE MINUTES

            Of the

COMMITTEE ON TECHNOLOGY

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

                    November 20, 2013
                    Start: 01:04 p.m.
                    Recess: 03:13 p.m.

HELD AT: 250 Broadway - Committee Rm,
                  14th Fl.

B E F O R E:
                  FERNANDO CABRERA
                  Chairperson

 COUNCIL MEMBERS:
                  GALE A. BREWER
                  LETITIA JAMES
                  G. OLIVER KOPPELL
                  MARK S. WEPRIN

1

2   [gavel]

3   CHAIRPERSON CABRERA: Good afternoon and

4 welcome to this hearing concerning DoITT.

5 Administration of local law 11, otherwise known as

6 open data law, open, open government data is the

7 concept that data that is generated by public

8 sector entities should be available to the public

9 electronically via the internet in open raw, raw

10 formats. Adopting open government data standards

11 promote inclusion of more individuals into the

12 governmental process by making access to

13 information easy and affordable. It also promotes

14 transparency and accountability by providing data

15 in its raw form enabling members of the public to

16 perform their own data analysis and draw their own

17 conclusions. In recent years the city has made an

18 enormous amount of public data available on

19 NYC[dot]gov. However much of this data was

20 available only in proprietary or visual formats

21 thus the usefulness of this data was limited. In

22 March 2012 the city enacted local law 11, otherwise

23 known as the open data law. Among other things the

24 law requires that within one year all public data

25 sets maintain by city agencies be made available on

the internet through a single web portal that is

linked to NYC[dot]gov and that the data sets be

made available in open formats in accordance with

certain technical standards published by DoITT. The

law also requires that DoITT submit a compliance

plan within 18 months that includes a summary

description of public data sets under the control

of each agency and when and how often they will be

published. DoITT released its compliance plan which

is, which it off, it refers to as the open data

plan this past September. While the city's

implementation of the open data law has been

impressive some issues remain including among other

things; the type of data, that are

postes[phonetic], posted on the city's open data

web portal, the dates that certain data will be

released, the way it's going to navigate on the web

portal itself. The committee looks forward to

hearing about implementation of local law 11 and

how DoITT is resolving these issues and now I'm

going to turn it over to the genesis of this bill

that was introduced and now law council member

Brewer.

1

2   COUNCIL MEMBER BREWER: Thank you very

3   much council Cabrera and I also want to thank many

4   of the, many of the people who worked on this;

5   certainly Colin Howell who's council to the

6   committee, Jeff Baker, and Jon Auerbach who are

7   also a part of the team in the speaker's office

8   working on this and Will Colgrove from my office

9   and the amazing transparency group which is the

10 folks coming together from good government and from

11 technology. And you should just know that it's such

12 an exciting topic that in the transition tent

13 tonight from six to eight there's going to be even

14 more discussion of open data. So it's something

15 that is on everyone's mind and I certainly want to

16 thank Mike Flowers and commissioner Merchant and

17 the many people who have made the open data portal

18 possible. And I always thank Carol Post who was a

19 former commissioner of DoITT and who was

20 instrumental. It is one of my proudest

21 accomplishments as a legislator and I think, thanks

22 to all the people I mentioned, New York City is at

23 the forefront of the government transparency

24 movement around the world. And I'm just delighted

25 that all of you, because it is a joint process have

this possible. I think to date we have about 11

hundred distinct data sets from over 60 agencies on

the open portal data. And the plan which the chair

mentioned which was released in September of this

year contains another 434 data sets identified by

city agencies as being subject to the open data

law. And I, I think that Mr. Flowers will talk more

about some others that may be available in January.

While this is positive we have serious questions

about some aspects of how the city is complying

with the open data plan. Advocates have identified

several important data sets that agencies maintain

which are not on the portal. And of serious concern

it's the fact that under technical standards

established by the city, all qualifying data sets

maintained by city agencies that are already on an

agency website and they were as I understand it

supposed to be posted to the open data portal by

March 7th of this year. And we believe that

compliance with this deadline is lacking and we

still do not have a comprehensive understanding of

a city's shortfall with regard to that deadline.

However I look forward to hearing from the city

about the plans and I want to thank again the staff

2 chairman Cabrera for this hearing today and we look

3 forward to not only to date but the, the follow-up

4 and the fact that this data can be implemented and

5 used is what I think our real goal is. Thank you

6 very much.

7   CHAIRPERSON CABRERA: Thank you so much

8 council member Brewer. JONATHAN AUERBACH.

9   JONATHAN AUERBACH: Alright thank you.

10 Thank you. [pause] So the, the idea here is to

11 graphically display some of the points that were

12 just raised. So I have here a few, a few slides. So

13 the number of state and local governments with open

14 data policies is growing. This graph with data

15 taken from the sunlight foundation shows the number

16 of policies adopted each year. Nearly 25 state and

17 local governments have open data policies, 17 being

18 adopted within the last two years. So for example

19 this is a bar chart here. Here you have Chicago and

20 Philadelphia that were created by executive order

21 between the years of 2012 and 2013. As you can see

22 from this chart the colors represent the legal

23 means by which these policies were adopted. Six of

24 the open data policies were created by executive

25 order while 15 were created by legislation similar

2   to New York City. So what is, what is open data?

3   Open data is raw, unaltered data released to the

4   public in a useable format via the open portal. How

5   is it different from regular data? It's all about

6   the presentation of the data. The ease in accuracy

7   with which users can manipulate the data and draw

8   their own conclusions. So as an example take 311

9   non-emergency service, 311 non-emergency service

10 request system. This slide is a screenshot of the

11 311 reporting tool from the New York City stat

12 website. It shows the 311 service request call

13 volume over the month of October and the fiscal

14 year to date. For example it tells us that the

15 Department of Finance had 16,426 calls through the

16 311 system that fell under the category of parking

17 ticket lookup here. This is an example of a

18 presentation where there is low control over the

19 data. That is, the data has been processed. The

20 data as presented here only answers questions

21 concerning the volume of calls each month. Weekly

22 calls for example would require an educated guess.

23 This slide is a screen shot of the NYC 311 map from

24 the city's 311 website. The tool shows us 311

25 service requests that are open or have been closed

2   within the last five days visually on a map. For

3   example we can see that on Broadway near 67th

4   Street right here there were several… [interpose]

5   [background comments]

6   JONATHAN AUERBACH: A district chosen at

7   random. There were several complaints of broken

8   meters. This presentation of the data gives us

9   medium control over the types of questions we can

10   ask. While we can filter service requests by

11   complaint, type, and location we cannot for example

12   aggregate the data and we could not reproduce the

13   results of the previous slide from this current

14   slide. And finally here is a screenshot of the open

15   portal. Like the map from the previous slide it

16   contains a filter feature and visualization tool

17   for examining the data. However most importantly

18   each individual service request is displayed in the

19   spreadsheet below. The data here can be downloaded

20   directly from the website of accessed via the API

21   and manipulated freely and therefore this

22   presentation of the data gives the user a high

23   level of control. I note highly control however

24   sometimes requires users to have a relatively high

25   degree of data sophistication. So the number of

2   data sets on the open portal has grown tremendously

3   according to a data set on the portal which

4   catalogs the open portals contents. The blue line

5   here on this side shows the growth of data sets on

6   the portal. However due to a feature of the portal

7   which allows users to generate new data sets much

8   of the data is redundant. The pink line below shows

9   the growth of datasets on the portal that have been

10   updated in the last month. From this graph we see

11   that roughly half the data sets on the open portal

12   have been updated in the last month. So here's

13   about a thousand and here we have slightly more

14   than 2,000. In addition to allowing users to

15   generate new data sets users can comment on

16   existing data sets. The portal contains 67 of these

17   comments. This slide shows the number of comments

18   each year that received a response. So we have

19   2011, 12, and 13. We can see here that 40 comments,

20   sorry that, roughly 40 comments were made last

21   year. About a quarter of these comments received a

22   response, probably more like a third. There are

23   more data sets on the portal than just 311 service

24   requests. And each data set is presented to the

25   users on the portal in a variety of ways. This

2   slide is a second example of the type of data sets

3   which can be found on the open portal.

4   The data set corresponding to this

5   screenshot contains information on the location of

6   Wi-Fi hotspots in New York City and some basic

7   descriptive information. As can be seen in the next

8   slide this data is presented to the portal, on the

9   portal to users as a map. On the next slide is a

10 third example of the type of data sets on the

11 portal. The data set corresponding to this

12 screenshot contains information on the number of

13 births by woman in New York City in 2007 in eight

14 different age groups. This data set is presented to

15 users as a pie chart. Unlike the 311 service

16 requests and Wi-Fi location data sets the

17 underlying information has been aggregated to make

18 the pie chart and cannot be filtered or broken down

19 below the eight groups represented in order to

20 additional questions.

21   This slide was made using data from the

22 Department of Information Technology and

23 Telecommunications September 22nd open data plan.

24 The plan lists the anticipated release date of all

25 data sets covered by the open data law as

2 identified by each agency. This graph shows the

3 number of data sets to be released by quarter. As

4 can be seen a significant portion of the identified

5 data sets will be released in 2016 or before.

6 Another significant portion will be released in

7 2018, the latest time period allowed under the law.

8   And this last slide shows the number of

9 data sets to be released by the 15 largest

10 agencies. The green color represents the number of

11 data sets to be released before 2016 while the blue

12 color represents the number of data sets to be

13 released in the year 2018. This graph shows that

14 several agencies such as the Department of Finance

15 here and the School Construction Authority here

16 will release the majority of all data sets

17 identified in 2018. The office of the mayor for

18 example over here will release all identified data

19 sets before 2016.

20   [off mic comments]

21   COUNCIL MEMBER CABRERA: Going to turn

22 it over now to Mr. Mike Flowers from the office of

23 the mayor and then if you have anybody else who

24 going to be speaking they can introduce themselves.

25

2   Thank you so much it's all yours. Only if you like

3   to.

4   [pause]

5   MICHAEL FLOWERS: There we go. Alright.

6   Sounds good. Sorry about that. I'm going to start

7   again. Good afternoon my name is Michael Flowers.

8   I'm the Chief Analytics and Open Platform Officer

9   for the City of New York. And my role as the Chief

10 Open Platform Officer I am responsible for the

11 implementation of local law 11 the New York City

12 open data law. Here with me is my Chief of Staff

13 Nicholas O'Brien, Deputy Commissioner for

14 Application Development of the Department of

15 Information Technology and Telecommunications,

16 DoITT, Donald Sunderland DoITT's General Council

17 and DoITT's General Council Charles Fraser. My

18 office, the Mayor's Office of Data Analytics works

19 closely with DoITT to fulfill the requirements of

20 the open data law. The open data law was signed by

21 Mayor Michael Bloomberg in March 2012. The law was

22 the product of a collaborative effort including the

23 city's technology community, the nonprofit sector,

24 the council, city agencies, and the mayor's office.

25 The mayor has referred to the law as the most

2   ambitious and comprehensive open data legislation

3   in the country. It has served as the basis for a

4   number of other policies and initiatives throughout

5   the country and has the full support of the

6   administration. I'm here today to testify on the

7   progress on the implementation of the open data law

8   and highlight some of the successes as well as

9   challenges we've had since the law took effect. In

10   April Mayor Bloomberg and Chief Information and

11   Innovation Officer Rahul Merchant formally

12   appointed me Chief Open Platform Officer. Since

13   that time my office has focused relentlessly on

14   delivering high value data to New Yorkers. In July

15   the Department of City Planning released one of the

16   most sought after data sets PLUTO which contains

17   land use and geographic data at the tax level. In

18   October we released ACRIS property records and

19   parking ticket data from the Department of Finance.

20   We have also inclu[phonetic], focused on excuse me

21   automating data sets to ensure that the published

22   data are as up to date as possible. The primary

23   vehicle for delivering data to the New Yorkers, to

24   New Yorkers excuse me is the open data portal

25   accessible at NYC[dot]gov[backslash]data. We

2   currently have over 11 hundred unique sets

3   available up from 350 in 2011. 59 of those sets

4   have been fully automated and new automations are

5   coming online every month. In September we release

6   the inaugural New York City open data plan which

7   catalogues over 460 unreleased data sets covered

8   under the law and their planned release dates.

9   Earlier today we launched the beta version of the

10   open data dashboard to provide insight into the

11   progress of publishing the data sets included in

12   the open data plan. Currently we are in the process

13   of reviewing the lists submitted by mayoral

14   agencies pursuant to the open data plan to ensure

15   that all qualifying data is included, that it is

16   released at the earliest possible date, and that no

17   private data is inadvertently listed. The annual

18   update to the open data plan will be released to

19   the council on July 15th, 2014. In addition to

20   these activities which are mandated under the law

21   we have undertaken and are undertaking numerous

22   initiatives beyond what is required by the law to

23   increase transparency and facilitate the leveraging

24   of open data to provide meaningful insight. In May

25   we launched the New York City Developer Portal. The

2   Developer Portal is designed to provide data to a

3   more technical audience for the purpose of creating

4   commercial grade applications. The developer portal

5   includes application programming interfaces or APIs

6   from 311, Department of Transportation, Department

7   of Education, Housing Preservation Development

8   Comptroller, and the Department of City Planning.

9   This enables developers to build applications that

10 integrate directly with high value public data. In

11 June we announced the winners of the fourth annual

12 New York City BigApps Competition. The Competition

13 hosted by DoITT and the Economic Development

14 Corporation is designed to encourage creation of

15 online and mobile applications that make the data

16 we release more accessible to New Yorkers and to

17 jumpstart new business built on top of open data.

18 Since 2009 NYC BigApps has helped launch nearly 300

19 new applications. In September we updated the NYC

20 open data portal to improve the user experience and

21 to help New Yorkers navigate the massive amount of

22 data available. All data sets are categorized for

23 easy browsing and are fully searchable. We provide

24 background information on featured data sets to

25 increase awareness to the high value data sets on

2   the portal. We are currently featuring the newly

3   released parking ticket data. Information about how

4   to deal with parking tickets is one of the most

5   common information requests of 311 as we just saw.

6   This data will provide New Yorkers with greater

7   access and ability to address their tickets. We are

8   also using the open data portal to show

9   visualizations built with public data. These

10 visualizations make the information contained in

11 the data more accessible and understandable as you

12 just seek to understand their city. We are working

13 with agencies on an ongoing basis to enhance the

14 public's ability to access, consume, map, and

15 visualize New York City's data. By leveraging the

16 city's GIS expertise we have enabled users to

17 create different types of maps of more than 165

18 data sets through the open data portal including

19 hurricane evacuation zones, Wi-Fi hotspot

20 locations, parks, and building footprints. There

21 are many high value data sets that are generated by

22 vendors for delivery under contract to the city. We

23 believe to the greatest extent possible this data

24 should be released as open data. DoITT aggressively

25 negotiates for the intellectual property rights on

2   all data created, generated, or maintained by the

3   city's contractors and whenever possible works to

4   provide public access to that data. We are already

5   seeing the positive impact of the city's open data

6   efforts. For example the Department of Education is

7   working with Pediacities a former NYC BigApps

8   winner and the successful startup built on NYC open

9   data to launch a set of public APIs. Previously

10   middle school students applying to high school had

11   to slog through a phonebook sized catalogue or

12   navigate a very basic search page to get

13   information on DOE's nearly 700 high school

14   programs. For next year's application process

15   students will be able to access this information

16   through nationally recognized mobile and online

17   applications including some of the application,

18   excuse me some of the same applications that high

19   school seniors across the nation used to apply to

20   college. We're also starting to develop proofs of

21   concept around utilizing the open data portal for

22   agency operational needs. For example the Taxi and

23   Limousine Commission is directing fleet owners to

24   the automatically refreshed list of licensed taxi

25   drivers to verify that their driver's licenses are

2   current. Our hope is that by using the portal to

3   provide information to the public and other

4   agencies the city can save significant cost in new

5   IT projects to achieve the same ends.

6   One of the benefits of open data we are

7   working to quantify is the improvement in

8   fulfilling FOIA requests. During the month of

9   October EC, or the Environmental Control Board or

10 ECB received two FOIA requests that they were able

11 to meet by directing their requesters to the open

12 data portal. This not only makes it easier for the

13 agency to respond to requests but it also provides

14 faster, easier, and cheaper access to the data

15 potentially reducing the need for the public to

16 submit lengthy FOIA requests to gain access to the

17 information.

18   Another data set that was previously

19 only available via FOIA request or a fee is PLUTO

20 and map PLUTO, a mash up of land use and geographic

21 data from a range of agencies. This data has been

22 downloaded more than 579 times since it was

23 released at the end of July. In the year prior to

24 the release it had only been downloaded 141 times.

25 In the coming months we are focusing on releasing

2   automated feeds of newly available data including

3   season flu vaccination locations, farmer's markets,

4   Office of Emergency Management Incidents and

5   Notifications, and Office of Management and Budget

6   Revenue Expense and Capital Funds data. In parallel

7   we are working to automate existing data sets from

8   the Department of Housing Preservation, and

9   Development, Department of Transportation, and

10   Department of Environmental Protection just to

11   mention a few. We are also focusing on measuring

12   the economic benefit of open data. We are tracking

13   aggregate usage and have begun to catalogue the

14   applications built on top of the data. This is the

15   first step in our effort to rigorously quantify the

16   economic impact. These benefits come in many forms

17   ranging from agency efficiencies and cost avoidance

18   to the creation of new jobs and businesses made

19   possible by the release of the data. We are

20   confident that this analysis will demonstrate that

21   the city's investment in open data makes economic

22   sense and we look forward to presenting the result

23   of this effort the next time we have the

24   opportunity to testify.

25

2   I would like to thank the council

3   Technology Committee for providing me the

4   opportunity to testify today and we are happy to

5   answer any questions you may have.

6   CHAIRPERSON CABRERA: I just got a

7   couple of questions and then I'm going to turn it

8   over to my colleagues. The first question I had,

9   this is in terms of compliance, many agencies miss

10 their March 2013 deadline for publishing data sets

11 that were already on their websites. Can you

12 explain why this was the case and whether all of

13 these data sets have been added to the open portal?

14   MICHAEL FLOWERS: So as, as a threshold

15 matter you're right. They missed them. And when

16 they missed them, I missed them because it's my job

17 to ensure compliance. What I can tell you is before

18 I turn it over to commissioner Sunderland is that

19 it's a big city.

20   [laughter]

21   MICHAEL FLOWERS: We are aggressively

22 working with every agency and what is there 40 some

23 odd.

24   [background comments]

25

2   MICHAEL FLOWERS: Right 45 and there are

3   a multitude of data sets that could be covered,

4   might be covered. I think there was a bar chart

5   presented by that gentleman. It actually is a

6   pretty good map as to where agencies are in

7   determining what, what data should and shouldn't be

8   released. So there's insight to be gleamed from

9   that right? And you know the fact of the matter is

10 that we, we see that document in many ways as

11 organic. I understand there's a deadline in there

12 and we try to meet those deadlines wherever

13 possible but we'd rather be right. And what we want

14 to do is continuously see that it's something that

15 we're obligated to constantly update whenever

16 possible. The dashboard that just kind of got put

17 up will help us do that in a more dynamic manor

18 than putting it on paper and then submitting a

19 report but please go ahead sir.

20   COMMISSIONER SUNDERLAND: [off mic] Sure

21 I mean basic, I mean we agree that we… [interpose]

22   CHAIRPERSON CABRERA: I'm sorry if you

23 could introduce yourself… [interpose]

24   COMMISSIONER SUNDERLAND: [off mic] I'm

25 sorry.

2   CHAIRPERSON CABRERA: …and if you could

3   pull the mic closer to you.

4   COMMISSIONER SUNDERLAND: Sure.

5   CHAIRPERSON CABRERA: Thank you.

6   COMMISSIONER SUNDERLAND: I'm Don

7   Sunderland. I'm the Deputy Commissioner for

8   Application Development for DoITT. We estimate that

9   about 70 percent of the data that, that should have

10 been covered by March had, was covered and we're

11 working on getting the rest. In many instances I

12 mean the, the agencies have deferred as Mike Said

13 we've pursued them vigorously, Mike in particular

14 has pursued them vigorously. In many cases they,

15 they cited Sandy as being a reason why they were

16 slowed down and much of the work that they

17 delivered… [interpose]

18   JONATHAN AUERBACH: Reasonably.

19   COMMISSIONER SUNDERLAND: Yeah

20 reasonably.

21   MICHAEL FLOWERS: I would say.

22   COMMISSIONER SUNDERLAND: Right. Right.

23 All that notwithstanding I'd like to reiterate his

24 point which is basically this is an ongoing work in

25 progress. Not only in delivering those sets but in

identifying new sets that would fall under the

purview of the law. So we really envision our

mission at this point as, as continuing to

vigorously pursue the sets, identifying new sets as

they arise and, and get them into the most usable

form possible. We are, we're, we're deeply

committed to that.

CHAIRPERSON CABRERA: Do you have an

estimate as to when we will reach 100 percent…

[interpose]

COMMISSIONER SUNDERLAND: We actually…

[interpose]

CHAIRPERSON CABRERA: …or close to…

COMMISSIONER SUNDERLAND: You want to… I

mean I don't think you ever reach 100 percent. Our,

our systems are constantly changing right. And

point a fact you had the Department of Health and

the Department of Consumer Affairs recently change

their background systems. That generates different

kinds of data which is good.

CHAIRPERSON CABRERA: Mm-hmm.

COMMISSIONER SUNDERLAND: We like it a

lot. But those constant changes basically are

always going to move the goal line ahead of us.

2   Right? So I don't think we'll ever reach the

3   platonic ideal of 100 percent compliance but rather

4   we are 100 percent committed to continuously trying

5   to ensure compliance.

6   CHAIRPERSON CABRERA: What would be an

7   acceptable percentage for you to, to be in, in a

8   current state? If you are at 90 percent, 95

9   percent? I, I, I, I used… What would be an

10 acceptable… [interpose]

11   COMMISSIONER SUNDERLAND: I'm not trying

12 to quivel[phonetic] on this. I, I'm uncomfortable

13 putting a number on it right? I guess what I'm

14 trying to say is I'm not comfortable with anything

15 less than 100 percent.

16   CHAIRPERSON CABRERA: Gotcha.

17   COMMISSIONER SUNDERLAND: So we'll

18 constantly have to work towards that. We have a

19 constantly shifting landscape, and a constantly

20 shifting set of actors. So as those different

21 moving pieces align the goal always remains. We

22 have to get this out.

23

24   CHAIRPERSON CABRERA: So just help me

25 understand here. Back in March what percentage?

2   COMMISSIONER SUNDERLAND: [off mic] 70.

3   CHAIRPERSON CABRERA: 70, 70 percent

4   that what? That we reach or we didn't reach?

5   [off mic]

6   COMMISSIONER SUNDERLAND: [off mic] 70

7   percent reached.

8   CHAIRPERSON CABRERA: Reach. And what is

9   it now? What are, where are you at now?

10   [background comment]

11   COMMISSIONER SUNDERLAND: [off mic] We'd

12   have to reassess it at this point. I'm sorry.

13   COMMISSIONER SUNDERLAND: We'd have to

14   reassess it. It's ongoing workforce at this point.

15   I mean we intend to have a, a thorough breakdown of

16   it in the July plan. When we… [crosstalk]

17   [interpose]

18   CHAIRPERSON CABRERA: Okay.

19   COMMISSIONER SUNDERLAND: …and we'll

20   give you a complete update at that point.

21   CHAIRPERSON CABREARA: But since March

22   have you checked?

23   MICHAEL FLOWERS: Well we keep adding

24   but we don't, we don't, we don't… [interpose]

25   CHAIRPERSON CABRERA: Okay.

2   MICHAEL FLOWERS: … we don't recomputed

3   every month. I mean,… [interpose]

4   CHAIRPERSON CABRERA: Okay.

5   MICHAEL FLOWERS: …and, and, and maybe

6   that's, that's a function that could go into the

7   dashboard. I mean maybe that's, that's a legitimate

8   function to be in there and we acknowledge that

9   that probably belongs.

10   CHAIRPERSON CABRERA: I just got one

11   more question and then I'm going to turn it over to

12   council member Brewer. How do you determine whether

13   an agency is compliant with the law? Is there a

14   formal process?

15   COMMISSIONER SUNDERLAND: I mean you

16   know I'd like to turn this over to general council.

17   CHARLES FRASER: I'm Charles Fraser. I'm

18   the General Council at DoITT. I think I would

19   answer that in two parts. One is the technical

20   people, the IT people at DoITT are aware to some

21   extent of what agencies have their websites because

22   they can, they can look at the websites. And so

23   they can tell if we've posted those through the

24   open data portal. Beyond that what public data sets

25   that an agency has that are not on the website we,

2 we have no way to know externally. We can only rely

3 on them and so I'm answering as general council

4 because I worked very much with the agency general

5 councils around the city to get them to identify

6 their public data sets for us so that our IT people

7 could work with their IT people to get them posted

8 or get them listed in the plan. So we're, we're

9 very much dependent on the agencies to tell us what

10 they have unless it's already on the website. I, I

11 will say one other qualification. When I say we're

12 very much dependent on the agencies; if the

13 community tells us they know of a database that's a

14 public data set and, and it's neither in the plan

15 nor on the portal we, we will follow-up with the

16 agency general councils on asking them about it.

17   MICHAEL FLOWERS: I'd say we like to

18 correct one of my earlier answers too. The 70

19 percent is to date.

20   CHAIRPERSON CABRERA: Okay.

21   MICHAEL FLOWERS: Though it's not part

22 of the dashboard at this point but it's a good

23 point that it should become part of the dashboard

24 in ongoing… [off mic]

25

COMMITTEE ON TECHNOLOGY   29

2   CHAIRPERSON CABREARA: Okay. Thank you.

3   Thank you so much. I'm going to turn it over to

4   council member Brewer.

5   COUNCIL MEMBER BREWER: Thank you very

6   much. I just, one thing that, not necessarily tech

7   but Mike when you mentioned the issue of you know

8   having a lot material online which I obviously

9   support, do we have like a community technology

10   centers which Time Warner promised during the

11   franchise agreements around the city so people… Or,

12   or is it they have to go to the library? What are

13   we doing to make sure the city is tech savvy so

14   that people can access all this information. And

15   that may seem like a strange question from me but I

16   am supportive of people being educated so they can

17   access.

18   MICHAEL FLOWERS: I don't think that's a

19   strange question at all because it, it really kind

20   of covers… You know we don't do open data for the

21   sake of open data right?

22   COUNCIL MEMBER BREWER: Well I know

23   that.

24   [laughter]

25

1

2   MICHAEL FLOWERS: We, we, we just don't

3   have the money to…

4   COUNCIL MEMBER BREWER: Okay.

5   MICHAEL FLOWERS: …to allocate…

6   [interpose]

7   COUNCIL MEMBER BREWER: Okay.

8   MICHAEL FLOWERS: …towards that…

9   [interpose]

10   COUNCIL MEMBER BREWER: But what are we

11 doing to… [interpose]

12   [crosstalk]

13   MICHAEL FLOWERS: …ideal. So…

14 [interpose]

15   COUNCIL MEMBER BREWER: What are we

16 doing on the CTC?

17   MICHAEL FLOWERS: …so what we're trying

18 to do in a lot of different fronts is you know

19 library's one of them.

20   COUNCIL MEMBER BREWER: Yeah.

21   MICHAEL FLOWERS: Right? We want to…

22 [crosstalk] [interpose]

23   COUNCIL MEMBER BREWER: They're hardly

24 open though. Go ahead.

25

1

2   MICHAEL FLOWERS: …you know. Correct.

3   Part of it, part of the response. Right? I don't

4   think there's a single silver bullet that will get

5   us to the stage where we are converting our open

6   data into insight right. What we try and do is use

7   a multitude of the city's resources. One of those

8   is the open data portal. Others of those are, you

9   know, the interaction with the nonprofit community,

10 the entrepreneurial community, and the academic

11 community who tells us you're falling down on this.

12   COUNCIL MEMBER BREWER: Okay.

13   MICHAEL FLOWERS: Let's do better. So…

14   COUNCIL MEMBER BREWER: Alright. I, I

15 won't belabor it. I'm saying Time Warner is

16 supposed to come up with 100 CTCs. I don't know

17 where they are. Somebody needs to tell us.

18   MICHAEL FLOWERS: Okay. They, they've

19 launched 10 to date and they're current…

20 [interpose]

21   COUNCIL MEMBER BREWER: 10?

22   MICHAEL FLOWERS: 10 to date and there

23 are 30 more coming in the, in the near future.

24   COUNCIL MEMBER BREWER: So it's only,

25 it's only 40, not 100?

2   MICHAEL FLOWERS: That's how far we've

3   gotten so far.

4   COUNCIL MEMBER BREWER: Okay. Well I

5   thought it was a hundred and we need, I think the

6   city council needs to know where they are.

7   MICHAEL FLOWERS: Okay.

8   COUNCIL MEMBER BREWER: Maybe that's a

9   data set. Is it?

10   MICHAEL FLOWERS: [laughter] I don't

11   believe it's. Maybe it's another dashboard number.

12   COUNCIL MEMBER BREWER: Okay. Number two

13   is I know that there was a September plan which we

14   all have, I, I read it, I can't, It's hard to

15   sometimes make sure that what's on the September

16   plan is also on the portal. So my question is how

17   do you… Is this an accurate summary of all the data

18   sets on the portal? In other words do you check to

19   see if this matches up? Again this is a lot of

20   retail work.

21   NICHOLAS O'BRIAN: My name's Nicholas

22   O'Brian. I'm the Chief of Staff, Mayor's Office of

23   Data Analytics. The plan represents everything we

24   plan to publish in the future. It does not cover

25   things that are already on the portal. The progress

2   on the plan is on the dashboard which was released

3   today and is available online.

4   COUNCIL MEMBER BREWER: I'm glad we have

5   the hearing. Thank you. The other question I have

6   is what's the updating? I know that you indicated

7   that it is some point, some part reliant on what I

8   would call complaint or positive complaint or

9   complaint based system. In other words this isn't

10 there. This isn't correct. But how do you plan to

11 do the update? And how do you, I think there was, I

12 don't know if it was a slide but there was a

13 discussion about how many questions have actually

14 been answered. You getting a lot of feedback on the

15 site which is good but are you able to answer them?

16 How do you answer them? And how do you follow-up

17 when people have issues that they think are

18 corrected? So it's an update question… [interpose]

19   NICHOLAS O'BRIAN: Yeah.

20   COUNCIL MEMBER BREWER: …and how do you…

21 'Cause I remember when we had the many many many

22 many meetings on doing the data and passing the

23 bill and the big issue was how do you keep it

24 updated and the answer was complaint based system.

25

2   So if that's the answer how are we making sure that

3   their complaints are answered or the suggestions?

4   NICHOLAS O'BRIAN: I, I can answer the

5   first part of your question about how, which I

6   think you directed to how we're going to do the

7   July update and… [interpose]

8   COUNCIL MEMBER BREWER: Yes.

9   NICHOLAS O'BRIAN: …thereafter? I think

10 the biggest thing we learned in the first round of

11 this, the September plan was how complicated it is,

12 how, how challenging it is. And so we have planned

13 to convene our, our working group, those of us who

14 may remain employed in January to begin planning

15 and begin working on the update that will be due in

16 July that working group included Mr. Flowers'

17 office, my office, the law department, and we have

18 to have people from my, from DoITT. And, and we

19 reached out through general councils as well as the

20 IT staffs of the various agencies. So we intend to

21 do a sort of a full press, almost a redo of the

22 plan for 2014 because we, we know for sure, I mean

23 just given the complexity of it, the number we're

24 talking about, we know we missed some even before

25 anyone comes in and tells us; hey I know about this

2   data set that you missed, just do to the complexity

3   of it. So we're going to do a, as I say, almost a

4   full redo starting in January for, for July.

5   MICHAEL FLOWERS: On the open data

6   portal and as part of the developer community we do

7   have the ability to register comments and we engage

8   in conversations with developers and data

9   consumers. We also take conventional means. I've

10 actually gotten personal calls that I've passed on

11 and started… [crosstalk] [interpose]

12   COUNCIL MEMBER BREWER: I'm sure with

13 this community you have.

14   [laughter]

15   MICHAEL FLOWERS: So, so it's, it's in

16 the, it's, it's a significantly responsive

17 community. It's a good community to work with and

18 we try to be… [interpose]

19   COUNCIL MEMBER BREWER: So do you think

20 that you're able to answer… I mean I know I have a

21 huge email list and I answer emails. I am familiar

22 with trying to be up to date with answering but do

23 you think… 'Cause we had a, a graph, I'm not saying

24 it's correct… [interpose]

25   MICHAEL FLOWERS: Yeah.

2   COUNCIL MEMBER BREWER: …that said a lot

3   of them were not answered. I don't know how one

4   would determine… [interpose]

5   MICHAEL FLOWERS: Yeah.

6   COUNCIL MEMBER BREWER: …whether they

7   are or they are not. How are you following that?

8   'Cause it was part of the reason that we, I mean it

9   was a way in which we're gaging material. It's not

10   like a frivolous system in order… You know the

11   system was put in. The complaint system if you call

12   it that… [interpose]

13   MICHAEL FLOWERS: Yeah.

14   COUNCIL MEMBER BREWER: …to be able to

15   keep it up, keep the, the sources updated 'because

16   you know you can't as a city agency keep all that

17   information and know it all. That's why the… So how

18   are you monitoring the responses?

19   MICHAEL FLOWERS: Okay at this point I

20   will have to come back to you on how it is that

21   we're following up on that…

22   COUNCIL MEMBER BREWER: Okay.

23   MICHAEL FLOWERS: …how we're tracking

24   it.

25   COUNCIL MEMBER BREWER: Alright.

1

2   MICHAEL FLOWERS: I don't have that…

3   [crosstalk]

4   COUNCIL MEMBER BREWER: I appreciate it.

5   The other issue in that tracking is I know the

6   public has certain data sets that they're

7   interested in. You could imagine which ones they

8   are. How do you track if they're, again, this is

9   the same kind of question to know which agencies

10 have dragged their feet on releasing data that they

11 know the public is interested in. Is that something

12 that you're able to at least think about if you

13 can't track it specifically.

14   MICHAEL FLOWERS: It's, it's, it's

15 actually something that we're integrally familiar

16 with because we end up being the middle person and

17 we…

18   COUNCIL MEMBER BREWER: Right.

19   MICHAEL FLOWERS: …receive all the

20 pressure that's coming in and all the… And we, we

21 try to transfer that pressure as best we can to the

22 agency who's engaged and there's several in which

23 we, we're continued on continuous improvement on

24 those issues.

25

2   NICHOLAS O'BRIAN: Yeah and if I could

3   just add to that. You know I mean that's the whole

4   point of having my position out at the mayor's

5   office as opposed to you know DoITT going agency to

6   agency. So, I mean there's, there's a constant

7   dialogue on that which we want to get out as

8   quickly as possible. You have to prioritize and

9   when I don't want to say the agencies have been

10 recalcitrant or not recalcitrant it's just a matter

11 of capacity. And to the extent that we sit there

12 and say okay well we'll throw some, some labor your

13 way to get this out more quickly it's really based

14 as much on public input as it is… You know I noted

15 as, as an example of the benefits of the open data

16 is, is internal efficiency.

17   COUNCIL MEMBER BREWER: Right.

18   NICHOLAS O'BRIAN: So I use that data.

19   COUNCIL MEMBER BREWER: But in the

20 Office of Operations is excited about it also.

21   NICHOLAS O'BRIAN: You know. Yeah and so

22 I mean, there are plenty of, plenty of city

23 agencies that are now looking to each other's data

24 through the portal as opposed to you know kind of

25 making a phone call. So those priorities are set by

2   a number of different values. Not you know least

3   among them is, is, is public interest.

4   COUNCIL MEMBER BREWER: Okay. I mean I

5   hope at some point you will track the requests for

6   information, complaints, whichever you want to call

7   it because that is part of why we passed the bill.

8   It was our only safety net in terms of compliance.

9   I mean.

10   NICHOLAS O'BRIAN: That'll be our first

11 step… [crosstalk] [interpose]

12   COUNCIL MEMBER BREWER: You can't fine

13 an agency right? So if the issue is a complaint and

14 the fine is the public.

15   NICHOLAS O'BRIAN: Yep.

16   COUNCIL MEMBER BREWER: A hundred

17 meetings on that topic.

18   [laughter]

19   COUNCIL MEMBER BREWER: It's also our

20 understanding that you have a compiled a list of

21 data sets from agency websites. I think that

22 belittle[phonetic] bit the chairman asked this

23 question but do you have such a list? And do you

24 plan to make it available? In other words we know

25 that you're trying do what the chairman was

2   suggesting in terms of what's out there. Do you

3   have such a list and is it going to be like

4   available in maybe January? Or what's going on with

5   this list?

6   MICHAEL FLOWERS: So we have reviewed

7   just as part of the process everything that is

8   available on NYC[dot]gov. We are providing that to

9   the agencies for their review and, and are asking

10 them if they do believe that it is a qualifying

11 data set that it should be placed on the open data

12 plan by the update for July 15th.

13   COUNCIL MEMBER BREWER: Okay. And then

14 the plan as you know lists 434 data sets to be

15 released by agencies by 2018. We hated that date

16 but we live with it. So is this a complete universe

17 pursuant to the, to the law or are there other data

18 sets you know such as agency performance data

19 collected by the city and there are parts of the

20 law that require that data sets that cannot be

21 available be identified and indicated why. I mean

22 none of us want our, I, I understand when we passed

23 the law we were following pretty much the FOIL

24 discussion. In other words what's available under

25 FOIL should be available here. But I just didn't

2	know if there are… Are you working with agency to

3	say we cannot release this maybe because you know

4	the state doesn't let us or the federal government

5	doesn't let us or reasons why. Is that something

6	also that's part of your plan?

7	 MICHAEL FLOWERS: You mean to, to make

8	those reason… [interpose]

9	 COUNCIL MEMBER BREWER: Correct.

10	 MICHAEL FLOWERS: …public.

11	 COUNCIL MEMBER BREWER: In other words

12	my understanding is that's what we were part of a

13	discussion with the open data law is to indicate

14	why you cannot release something even though there

15	may be a data sate, data set.

16	 MICHAEL FLOWERS: I mean I think it's

17	certain, it, absolutely.

18	 COUNCIL MEMBER BREWER: Okay so when is

19	that all going to happen? Because 2018 we might be

20	dead by then?

21	 [laughter]

22	 COUNCIL MEMBER BREWER: We're all trying

23	to get this thing done. At 2018 a lot of agencies

24	in the compliance plan, they love to put in 2018,

25	2018, 2018, and I have the list right here.

1

2   MICHAEL FLOWERS: Well you know, it's a

3   threshold matter. I certainly hope we're all still

4   here but… [crosstalk] [interpose]

5   COUNCIL MEMBER BREWER: I do too.

6   [crosstalk] I'm just saying I want them done

7   sooner.

8   MICHAEL FLOWERS: When we reached out to

9   agencies we didn't know exactly what we were going

10   to get back. Now that we have seen the totality of

11   the plan we are working with agencies to ensure

12   that their technical capacity to get it out and

13   DoITT's technical capacity to add it to the portal

14   in an automatically refreshing manner align. So

15   part of the things that we're looking for in the

16   July 15th update is to smooth out that curb and

17   bring some of the ones, the data sets that are

18   listed for 2018 into a closer timeframe so that we,

19   and, and, and one that we know we can meet on the

20   technical end on, on both ends.

21   COUNCIL MEMBER BREWER: The other

22   question I have is do you find that, you mentioned

23   a couple of FOIL examples and I appreciate that

24   'because that was a reason to pass the bill. So are

25   you checking with all agencies to see if,

2   particularly those that get a lot of FOIL requests,

3   to see if you know their data matches up some of

4   their FOIL requests so we can actually implement

5   that idea.

6   MICHAEL FLOWERS: Yeah, I mean in

7   affect. Right? So it's not so much whether… There

8   are certain agencies like ECB that get maybe, you

9   know we've mentioned that there were two requests.

10 Those were actually two massive requests right?

11   COUNCIL MEMBER BREWER: Okay.

12   MICHAEL FLOWERS: And so the same holds

13 for, that is absolutely part of the dialogue when

14 I'm dealing with agencies. Right.

15   COUNCIL MEMBER BREWER: Does anybody

16 know how many agencies get what kind of FOIL

17 requests? I mean does that… [interpose]

18   MICHAEL FLOWERS: I don't know that off

19 the top of my head.

20   COUNCIL MEMBER BREWER: That would be a

21 great data set.

22   [laughter]

23   MICHAEL FLOWERS: The FOIL request data

24 set?

25

2   COUNCIL MEMBER BREWER: Yes. I mean some

3   people are crazy and they FOIL everybody all the

4   time. I know all those people.

5   [laughter]

6   COUNCIL MEMBER BREWER: But, but… They

7   probably live in my district.

8   [laughter]

9   COUNCIL MEMBER BREWER: But I also know

10 that people have legitimate reason. So I think that

11 would be a something to look at and then you could

12 compare that to where you're able to reduce that in

13 a year or so to say this is something that the law

14 was intended to do and it's a cost savings. So I,

15 I'm just saying that we did ask when the law was

16 passed, or during the discussions; could we get a

17 data set on the number of FOIL requests and xx. My

18 guess is PD gets a lot.

19   MICHAEL FLOWERS: It's, yeah, I don't

20 know.

21   COUNCIL MEMBER BREWER: I know.

22   [laughter]

23   MICHAEL FLOWERS: I mean I, I really

24 don't know but I think what you're saying is

25 perfectly reasonable… [interpose]

2   COUNCIL MEMBER BREWER: Alright so is

3   that something that… [interpose]

4   MICHAEL FLOWERS: …and makes sense.

5   COUNCIL MEMBER BREWER: …that we… All

6   I'm saying. So that's the FOIL and the issue of the

7   bill would be helpful. Do you have recommendations

8   to, for the future as to how to help agencies

9   automite[phonetic], automate data posting? Because

10   I know some of them have legacy systems or other

11   issues. Are they something… Do you have some

12   recommendations how to address how the data can be

13   automated for posting more regularly? That would

14   help with the updating.

15   MICHAEL FLOWERS: You know you want to

16   take that?

17   CHARLES FRASER: Yeah. You know at this

18   point it's, it's very much a function of the

19   technology state of the, of the systems of the

20   agency. In all cases without moving to new

21   platforms they're going to have to do some kind of

22   data extraction and that's always going to be a

23   manual process. What we try to do is provide them

24   with standards whereby we can automate the feed of

25

2   those files once they've been extracted from their

3   systems and we do have that in place.

4   COUNCIL MEMBER BREWER: And do we know

5   how many agencies have the kinds of, I don't know

6   if it's a Legacy, I hope it's not… I don't know

7   what kind of system it is that would make it more

8   challenging.

9   CHARLES FRASER: Well often times there

10   are things like access to data basis or, or even

11   manually kept records that have to be entered into

12   a, a spreadsheet to be set so there's a whole,

13   there's a whole panoply of approaches. And as we

14   identified the data there's no guarantee the data

15   is even associated with the system per say. So

16   there's a, there's a whole spread here. What we're

17   trying to do is create standards around the formats

18   for the data so as they do capture it it's in a

19   form in which it can be fed to us.

20   MICHAEL FLOWERS: And then just to kind

21   of add to that it's not just the technology

22   challenge there right? So the, the question that

23   becomes to what extent can we set up an incentive

24   structure right, where they want to do this anyway

25   and it eliminates some of the labor involved from a

2   technological standpoint to bring systems into

3   compliance. Right? And much of the work that we've

4   described in our, in our opening statement really

5   is about kind of changing the incentive structure

6   so that they want default to open so that when they

7   build these new systems they're defaulting you know

8   to a system that allows this data to go forward. So

9   it's a, it's a kind of a joint technological and

10 business process.

11   COUNCIL MEMBER BREWER: Okay one last

12 question because I know we want to hear from the

13 public. I'm just wondering, I, I'm very interested

14 in the comments from the public and could that be

15 posted in the data set in the future. Obviously

16 that would take some intelligence and commitment

17 from staff but that's something that would be I

18 think helpful for archives and interest in general.

19   UNKNOWN MALE: So if you visit the

20 nominate website you can see in, on a single

21 website exactly all of the data sets that had been

22 suggested. So it is in a unified view. We also take

23 comments on specific data sets. So if you go to a

24 particular data set, all the comments related to

25 that data set will be there. So we've organized it

2 in a way that we are looking both at comments on

3 data sets and on things that should be there and it

4 is organized and is available in a unified form.

5   COUNCIL MEMBER BREWER: Alright. Thank

6 you Mr. Chair. I could ask questions all day but

7 there are a lot of other people who'd like to

8 testify. Thank you.

9   CHAIRPERSON CABRERA: You can ask

10 questions all day. You have it like that. I have, I

11 did have one question here before I let you go. Two

12 quick things. One is; is the open data law helping

13 to reduce 311 reports?

14   UNKNOWN MALE: I don't know.

15   CHAIRPERSON CABRERA: You don't know?

16   UNKNOWN MALE: I mean, no I, I really

17 don't know that. We could seek to measure it. I

18 don't know, I don't know what you do to reduce 311

19 reports.

20   [laughter]

21   CHAIRPERSON CABRERA: Okay. Here's the

22 other thing. You know I was looking at the maps and

23 usually they have the little dots and then

24 sometimes it's like an overdose of dots. Is there

25 any way to like split the map? Like instead of the

2   whole outer Manhattan being in one map to split it

3   in two so you know… I think at one point it just

4   becomes one big blog. You know.

5   UNKNOWN MALE: I think density maps like

6   that, alright you're talking about the 311 map

7   right?

8   CHAIRPERSON CABRERA: Right.

9   UNKNOWN MALE: I mean density maps like

10 that are always prone to over representation of the

11 information right?

12   CHAIRPERSON CABRERA: Right.

13   UNKNOWN MALE: I mean we're constantly

14 working on it, figuring out better ways to

15 visualize that information. I think what you're

16 saying is that imminently reasonable.

17   CHAIRPERSON CABRERA: Okay. I want to

18 thank you all, great information and looking

19 forward to a better future with this open data law.

20 Thank you so much.

21   UNKNOWN MALE: Thank you very much.

22   CHAIRPERSON CABRERA: Now we're going to

23 have JONATHAN AUERBACH Kaehny from Reinvent Albany,

24 Noel Hidalgo Code for America, Rachael Fauss in

25 Citizen's Union and Gene Russianoff from NYPIRG.

2   [pause]

3   [background comments]

4   COUNCIL MEMBER BREWER: These are the

5   coolest people in New York. Just want, these ones

6   coming up.

7   [pause]

8   CHAIRPERSON CABRERA: You may begin

9   whoever wants to go first. Can introduce yourself.

10   JONATHAN AUERBACH KAEHNY: Hi my name is

11   JONATHAN AUERBACH Kaehny. I'm the Executive

12   Director of Reinvent Albany and I'm also co-chair

13   of the New York City Transparency Working Group.

14   Brought along today a copy of a report,

15   recommendations from all of our groups in the

16   transparency working group, it's right here, handed

17   this out to members and hope that you get a look at

18   that. That's a more extensive background than what

19   we're going to talk about today.

20   I have a couple basic points. One, all

21   of our groups strongly strongly support the open

22   data law and think it is extremely important for

23   the future of open government in New York City on

24   par with the Freedom of Information Law. This is a

25   foundational aspect of transparency and open

2 government in New York City and we're, we're just

3 scratching the surface of the importance of what

4 this law offers and the benefits that it offers.

5 And I think that's a very important point for the

6 incoming administration and city council who we

7 strongly urge to fully fund open data law staff and

8 make sure the implementation goes ahead full bore.

9 And we credit the Bloomberg administration and the

10 city council for really pushing the open data law.

11 I mean there's been a lot of bumps in the road but

12 council member Brewer's work and the work of Mike

13 Flowers has been extremely important. So we want to

14 give them a lot of credit up front for this.

15   A couple things. The open data law is

16 working in our view and it's working well. The

17 spirit of the law is being met by the

18 administration and by the council. And I say that

19 at the onset. There's very very important public

20 data sets that Mike Flowers mentioned that have

21 been released PLUTO, ACRIS, 311 complaints, things

22 that the public had been working literally decades

23 to get out. PLUTO, since the early 90s some have

24 been trying to get out. The journalists have been

25 regularly been siting the open data law; the New

2   York Times, The Guardian online, national public

3   radio. They're siting the flood maps. They're

4   siting other key data on from the Department of

5   Health regularly. So if you look at the info

6   graphics in their online publications you'll see

7   New York City open data portal credited in the

8   corner there. And that's a sign of success. That's

9   exactly what we want to see from the open data law.

10   There's apps developers that are using the open

11   data law to build lots of apps. We'll hear about

12   that which is a way of getting information out to

13   the public in lots of different channels. And

14   that's one of the basic ideas behind the open data

15   law. And as Mike Flowers mentioned government is

16   starting to use the open data law and in this case

17   the Taxi and Limousine Commission. But we know of

18   many many examples from agencies in which regular

19   agency workers are looking at the open data site in

20   the course of their everyday work. And this is

21   exactly what we want to see happen. Get this

22   valuable information out. So the law is working but

23   we think the open, the implementation does need

24   improvement. I commend the council for pressing the

25   administration on questions about why the deadline

2   on March 2013 was missed. All of the data that was

3   due on that is already on city websites in a closed

4   format. That's very important. So we know that data

5   exists. In fact we know it's online in a non-useful

6   format. So they should really be pressed on that

7   and that data should come out much much sooner than

8   2018. And we hope that at the next oversight

9   hearing that they're further pushed on that. In

10   fact I'd recommend this committee send a letter

11   asking them to push ahead with that because I, I

12   really don't see the excuse that they have since so

13   much of that data is in fact already in a

14   spreadsheet format or a digital format that allows

15   easy easy publishing in the open data format.

16   They're just, they're not good excuses for that.

17   I'll accept other excuses for many many other

18   things. The size of the city does matter but in

19   this particular case they've already got that

20   information on their own websites.

21   There's a couple other points. We see

22   overall a mismatch at this point. And this is part

23   of the, partly due to the fact that the law is very

24   new. But there's a mismatch between what the public

25   most wants to see and what data the city has made

2   available. And I'm sure many many people after me

3   will speak to specific data sets there. But much as

4   Freedom of Information Law is about pushing, it's

5   about struggle, it's about always trying to get

6   better. The open data law is also a struggle. It's

7   an ongoing push between the public and the

8   administration. And we are continuing to see many

9   agencies dragging their feet on data that they know

10 that the public is extremely interested in. For

11 instance NYPD crime and traffic crash data much

12 much requested. Most major American cities have

13 this data on their websites in a open data format

14 so apps developers can use and show. New York City

15 should do the same. That's the tip of the iceberg.

16 We mentioned about a hundred data sets that we

17 identify in the back of this report that we believe

18 should be published on an expedited basis. And

19 there're probably many more that others here after

20 me will mention as well. But the entire back of

21 this report are priority data sets that we think

22 are technically feasible and we've identified where

23 they are. Lastly we have a couple recommendations

24 for getting data out that the public is most

25 interested in. First, and council should be really

2   aware of this and continue to push on this, this is

3   part of what you could call an open information

4   ecosystem with 311, with FOIL, and with website

5   usage data from city websites which the city

6   collects using Google Analytics. All of that

7   information should inform what data sets come

8   first. If the city has limited resources it needs

9   to be picking the most important data sets first.

10 And we know we're getting tens of thousands of 311

11 requests, many many thousands of FOIL requests, and

12 we have millions of hits on city websites. So we

13 have, we know what the public is interested and the

14 data should be prioritized accordingly to the

15 public interest. The whole idea behind open data is

16 that the public can make apps, they can make

17 websites, they can use this data and spread this

18 information in ways that offload requests from 311

19 and from city websites and make it more easily

20 available. That's what this is all about.

21   Lastly we would suggest that council

22 strengthen this law by passing a new law which

23 we're calling 'One Strike, You're in' and that

24 would require that data that's released under a

25 FOIL request be automatically required to be

2   included in the city's open data portal. So for

3   instance right now there are numerous numerous FOIL

4   requests and sometimes lawsuits against major city

5   agencies in which data is released ad hock and then

6   it doesn't go into the city's open data portal. It

7   would make complete sense that once that's in, it's

8   in. So if council would like to strengthen this law

9   that's a way of using FOIL in combination with the

10   open data law to strengthen both. Thanks.

11   [background comments]

12   [pause]

13   RACHAEL FAUSS: Good afternoon chair

14   Cabrera and council member Brewer. My name is

15   Rachael Fauss. I'm the Policy and Research Manager

16   for Citizen's Union a nonpartisan good government

17   organization dedicated to making democracy work for

18   all New Yorkers. As you both know we've been

19   involved in the open data law process for a long

20   time. We've supported the bill when it was first

21   introduced and the law when it was passed. And we

22   welcome the city council's continued engagement on

23   this issue and as government become increasingly

24   data driven it's important that that data be pushed

25   out to the public. So before turning to our

2   evaluation of things that are missing and things

3   that should be done in terms of the open data law

4   we'd just like to reiterate some of the positive

5   effects that the law has had and why open data is

6   so important. I know JONATHAN AUERBACH alluded to

7   them and they're in the progress and challenges

8   report but just to, just to reiterate. High valued

9   data sets, many of them have already been put out.

10 There's widespread use of the open data portal.

11 Agency staff, public, academics, journalists, and I

12 would assume many council staffers as well. And

13 there's also been a lot of commercial use of the

14 portal. And I think you know that really shows that

15 it's the goal of open data in creating jobs and

16 fitting in with the tech economy is actually coming

17 to fruition as well. So in terms of next steps.

18 Something that we think is an important thing for

19 the council to do is actually lead by example in

20 this area and many of the council has put some data

21 sets in the open data portal but there's certain

22 things that we think could be added and would be

23 important. First off being city council expenditure

24 reports. This is something that actually the New

25 York State Legislature does. They're not often seen

2   as a model in this area but they release

3   expenditure reports that has information about

4   mailings, rent, costs of other things associated

5   with members, central staff, etcetera. And this

6   actually was supported by 31 incoming members of

7   the council in the Rules Reform Agenda. So this is

8   something that I know some of the members of the

9   council are thinking about but it's certainly

10 something that would be important to include on the

11 portal.

12   COUNCIL MEMBER BREWER: I agree except I

13 ask one incoming member to come to the open data

14 hearing and he said what's open data. Just so you

15 know.

16   [laughter]

17   RACHAEL FAUSS: Well that's a part of an

18 education process I think But…

19   [laughter]

20   RACHAEL FAUSS: Well I think you know on

21 that point you know I think members are interested

22 in learning information about each other. So if

23 more information about the council itself is online

24 the members themselves may learn about the value of

25 it.

2   COUNCIL MEMBER BREWER: You are being

3   very nice.

4   [laughter]

5   RACHAEL FAUSS: And another piece that

6   we'd ask I know that the compliance plan has data

7   from Leister the council's portal for, for its own

8   legislation, another piece of data that could be

9   included, should be included as part of that is the

10 legislative service requests or when a member

11 requests a bill be drafted. There is a list of that

12 that from what we understand but it's not made

13 available to the public currently. Another piece

14 that we think would be important for the council to

15 do is for agencies that aren't mayoral appointed

16 ones where the council actually makes appointments

17 or shares appointments with the mayor. We think you

18 could play an important role in ensuring some of

19 their data is put out to the public. For example

20 the city clerk has a city lobbyist database. They

21 haven't reported in the compliance plan for

22 example. And we think both they should report what

23 data they plan to publish on the portal but also

24 the city lobbying database only includes basic

25 information. It doesn't include many of the

2   underlying data that's reported by the public. And

3   another agency that was not included in the

4   compliance plan that should be is the campaign

5   finance board. The council also makes appointments

6   there. They have a very good database themselves

7   but there's other data other than expenditure and

8   campaign contribution information that they have

9   that would be important to put on the portal. For

10 the administration there's a few things that we'd

11 like to highlight as priority data sets. But

12 actually before that we agree that, and I know this

13 came up earlier in the questioning, there should be

14 a list of which agencies are subject to the open

15 data law that's made available to the public and to

16 the city council. This helps with oversight both on

17 your end but also on our end as a good government

18 organization. Though of course agencies that aren't

19 subject to the law can voluntarily disclose

20 information. I mentioned some agencies that haven't

21 submitted compliance plans; the Campaign Finance

22 Board and the City Clerk, there're few others that

23 weren't included that we think should be examined;

24 Board of Standards and Appeals, Civilian Complaint

25 Review Board, Commission to Combat Police

Corruption, and that's among others. We look at

agencies that are relevant to the work we do but

I'm sure there are more that perhaps weren't

included in the compliance plan. And then the

council I think should also look at other

legislative changes to strengthen the law. I know

JONATHAN AUERBACH mentioned the 'One Strike, You're

In', something that was discussed as part of the

bill. Negotiation was including a right of action

for the public for including data in there. As

we're seeing the rollout of open data that's

something we think that should be considered,

continued to be considered. And then as far as a

couple of priority data sets that we think should

be on a fast track, we've got them in the New York

City, the Transparency Working Group plan but just

one data I'd like to mention is the doing business

database. I think for a good government and

accountability it's a particularly important

database to release to the public through the

portal. So thank you.

GENE RUSSIANOFF: Good afternoon

chairman Cabrera and the member Brewer and staff.

I'm Gene Russianoff from Senior Attorney with the

2   New York Public Interest Research Group and along

3   with JONATHAN AUERBACH I'm co-chair of the

4   Transparency Working Group. For the main part of my

5   testimony I wanted to recount a story that I think

6   really conveys the difficulties that many of our

7   groups' NGOs have faced and tried to get that out

8   of the city. So my story starts in 1980 and was way

9   back when. And NYPIRG wanted to do a study of the

10   fairness of city property taxes. We were told by

11   the finance department that the data was not

12   available because the selling price of the homes

13   was taken off of tax stamps and that as tax

14   information they were confidential. We were, we

15   were about to give up when we discovered in the

16   catalogue for the city planning commission that

17   exact piece of data was for sale. So, but the sad

18   part of the story is that in our haste to get ahold

19   of this data we purchased the printout because

20   everything is a learning experience in life and as

21   a result of that we had to key punch nearly 23

22   thousand home sales and it's 30 years ago but I can

23   remember the feel of the cold metal at the IBM key

24   punch card and the giant stack of mistakes I had

25   made that I had to redo. But you know with dog it

2   persistence we turned out a report called city of

3   unequal neighbors that documented low income

4   homeowners were wacked with higher assessments than

5   was fair and we convinced the city of New York

6   under Mayor Edward Koch to reassess about 50

7   thousand homes. But it's still a painful

8   experience. Out of these experiences groups

9   convinced the 1989 New York City Charter Revision

10 Commission to create a Commission of Public

11 Information and Communication. And except for its

12 first year it went for 23 years without funding.

13 The Public Advocates Office which is a chair of

14 this commission is the public advocate you know

15 made real serious attempts to try and change that

16 reality but it wasn't. So my point here is that

17 it's been a severe challenge for our groups many

18 times to get data from the city. And the New York

19 City open data law should change that. The rest of

20 my testimony can be summed up in one word after

21 Rachael and JONATHAN AUERBACH's testimony which

22 would be ditto.

23   NOEL HIDALGO: Thanks Gene, honorable

24 council person and councilors and staff. It's a

25 pleasure to be in front of you this morning. I'm

Noel Hidalgo. I'm a civic technologist and a civic

hacker. I'm the former Director of Technology

Innovation at the New York State Senate the

cofounder of Beta NYC. We're building a connected

New York City by the people for the people for the

21st Century and we do this through weekly

meetings. And in this room you'll actually find a

number of us. We are hackers, yackers, and mappers

and we are grateful for the city's data catalogue

and we're here to outline what we need to make our

communities safer and smarter. So back in 2009 we

started meeting and we were primarily a community

of civic technologists. We're really passionate

about being the technology vanguard. Over the last

few years we've partnered with several city

agencies to explore how data and tools can improve

the lives of New Yorkers. We're members of the NYC

Digital Code Core which is supposed to be the

National Guard for nerds here in the city. In 2013

we hosted and worked with NYC DoITT, NYC Digital,

NYC DOT, NYC Financial Empowerment Centers, the

EDC, Citi Bike, and NYU Wagner School just to name

a few of the organizations that we're working with.

Because of your leadership NYC is now seen as one

2 of the premier cities for civic hacking. Local law

3 11 of 2012 ushered in a new era of transparency and

4 opportunity and now I will outline three areas;

5 wins, frustrations, and opportunities.

6   One of the biggest win is the fact that

7 the New York City civic hacker community has grown

8 110 percent over the last year. We started the year

9 with 600 members and we're now a 13 hundred people

10 who are meeting on a weekly basis. We've done this

11 through 43 events. We pretty much meet as I said

12 weekly. In the spring we worked with the EDC to

13 reframe the BigApps through advocacy we got them to

14 focus on building communities and companies

15 'because we see that data is an opportunity for

16 this. So out of the BigApps seven apps won 125

17 thousand dollars. 42 data providers offered their

18 data. 517 individuals participated. 120 projects

19 were created. And 54 were eligible to win. This is

20 a vast ecosystem to what JONATHAN AUERBACH was

21 talking about a little earlier. The other big win

22 was PLUTO and ACRIS. The liberation of PLUTO and

23 ACRIS has fueled an explosive demand to understand

24 the city's built environment and property records.

25 StreetEasy, a made in NYC company and one of our

2   hack night partners was acquired this year for 50

3   million dollars. Their liber, the liberation of

4   these two data sets will produce more of these

5   success stories.

6   Some of our frustrations. Primarily

7   it's around NYPD and Citi Bike. While they are

8   independent of each other they have provided some

9   of the most rewarding conversations but also

10 frustrating. We have a monthly hack night with NYU

11 Rudin where 30 applications, public safety, and

12 bike apps have been created but because of the poor

13 data formatting we're forced to scrape the data.

14 This is a process that decreases the data's quality

15 and goes against the NYPD's testimony of usability

16 and accuracy. We want safer and more accessible. We

17 want a safer and more accessible city. And if we

18 have the data we could help. Some of the other

19 frustrations is around machine readability. We

20 continue to see data being locked away in PDFs or

21 'saved as' in air quotes as spread sheets that were

22 formatted to be printed and read by humans. When

23 they're released many of these data sets are

24 aggregated so they're, they're synthesized for us

25 as if we're not intelligent enough to read the pure

2 data. And what we want as civic hackers is that

3 data to be disaggregated and frequently updated.

4 From our research if a report or data is hosted

5 away from the NYC data catalogue and on an agency's

6 site there is a higher likelihood of that data not

7 being machine readable and in our mind that

8 violates the spirit of the law. I, in my testimony

9 have two examples. One of them is NYPD where we had

10 a, we met in front of the Transportation Committee

11 last month and gave explicit testimony. In summary

12 they said we already have hackers working on this

13 data. The reality is we're the hackers and we're

14 frustrated with the data and we want it to be

15 better and they said no. The other example is

16 Department of Transportation. They produce a number

17 of feeds, data feeds that are human readable but we

18 can't ingest them into our system and if we had

19 this data we would be able to build better

20 notification tools for drivers and cyclists. Some

21 of our other frustrations were enumerated at the

22 very beginning around the timeliness of the open

23 data plan. Only 75 of the data sets, 75 out of 434

24 data sets are going to be updated on a daily or

25 weekly basis? And that's a frustration for us

2   because we want to build useful apps and timeliness

3   of data release helps us build useful and friendly

4   apps. Example, New York City Schools, said almost

5   all of the data that they're going to be releasing

6   is soon to be released between now and the 31st of

7   December 2018 even though they just produced this

8   really awesome data set for high school students we

9   still don't know when the rest of their data is

10   going to be coming out. And the Department of

11   Mental Hygiene, Health and Mental Hygiene has data

12   sets that are coming out in 2016, 2017, and 2018.

13   We don't actually… These are critical health

14   inspection data but we don't actually know when or

15   how they're going to be releasing them.

16   So improvements, specific improvements

17   to the NYC open data law. One is to improve the

18   scope of the law. We heard beforehand state

19   chartered agencies, franchise agreements, and

20   county data sets, for example Richmond County their

21   property sales and court records are file under

22   county law so therefore they are not included in

23   the data, in the city data mine. And additionally

24   we completely support the One Strike and You're In

25   policy, machine readability, disaggregation of

2   data, and preferably error reporting. This says

3   it's something that hasn't come up before but the

4   law states that we want machine readability. We

5   completely agree it's really hard to get those

6   agencies to comply with that. The disaggregate data

7   is something that we're advocating for, we really

8   really want, it's been hard to get. And then

9   finally is around error reporting we'd like to see

10   some type of policy or legal construct around our

11   ability, the peoples' ability to report back

12   inconsistent data or issues of data quality

13   control. We would like to improve the frequency of

14   the data publication as I previously stated and

15   publish the data in common data standards. In 2012

16   Code for America, Yelp[dot]com, the city of San

17   Francisco, and New York City worked together to

18   produce a commonly accessible file format called

19   LIVES, Local Inspector Value Entry Specification.

20   It's a data standard for health inspection scores

21   essentially the A, B, C, and Ds that we see all

22   around in the restaurants. Sometime last year New

23   York City pulled out. We're not sure exactly why

24   but the current NYC restaurant inspection score

25   database is overly complicated and continues to be

COMMITTEE ON TECHNOLOGY

2   filled with inconsistencies and if we could get the

3   city to adopt more of these standards we would be

4   able to quickly turn around applications for you

5   that improve the lives of all New Yorkers. So thank

6   you for your time.

7   CHAIRPERSON CABRERA: Thank you so much.

8   Quick question before I turn it over to council

9   member Brewer. Was there anything that

10 administration, you heard them speak, comments that

11 frustrated you, concerned you, any red flags.

12   JONATHAN AUERBACH KAEHNY: [off mic]

13 Well I would…

14   JONATHAN AUERBACH KAEHNY: First I mean

15 we think that the mayor's office and Mike Flowers,

16 and Nick O'Brian have been working really really

17 hard and they're not going to throw any agencies

18 under the bus. They're, it's very clear however

19 that many agencies are dragging their feet on a

20 compliance and basically just refusing to, refusing

21 requests of DoITT and the Mayor's Office on this.

22 And we hear, you know firsthand account so we know

23 it's happening, we, we mentioned a couple agencies;

24 NYPD on crime and traffic, City DOT on Citi Bike

25 data Noel Mentioned but it also comes with DOE and

2   DEP and All the big main line agencies. They have a

3   lot of data sets that are of huge public interest

4   that they're really dragging on. And in some cases

5   it's technical issues with just the scale and the

6   reliability of the data. But in other cases it's

7   because they want control and they want control

8   over the release. So you know that's something

9   that's hard for them to say and go through. The,

10 it's very very important that the data dashboard be

11 effective because as Gale noted in her persistent

12 questioning we really don't have a lot of leverage

13 on the, we the public. We can't sue agencies to

14 comply with the open data law. There is no public

15 right of action on this. So public complaint and

16 public protest which has to be informed by the data

17 dashboard is very very important. And the, the

18 other piece though that we have in terms of

19 compliance and accountability are oversight

20 hearings like this one. So your questioning is

21 super important for the future success and

22 viability of this law. But I would point to basic

23 agency noncompliance as remaining a big problem. I

24 mean this is just like FOIL. I mentioned that in

25

2 that some agencies simply do not want information

3 easily available to the public.

4   CHAIRPERSON CABRERA: Thank you so much.

5 And I want to recognize council member

6 Tish[phonetic] that is has join us. Thank you so

7 much future of public advocate. Council member

8 Brewer.

9   COUNCIL MEMBER BREWER: Thank you very

10 much for all your efforts on behalf of

11 Transparency. I think you add it all up it's a

12 awful lot of time and commitment. I, I just have a

13 couple of questions. First of all did we have any

14 discussions and maybe I'm wrong with some kind of

15 an advisory board to the agencies or is that make

16 any sense? Otherwise you know the hearings make

17 sense, the comments that are organized in the

18 fashion it was described but I think you probably

19 need even more in talking about the dashboard. So

20 you know it's something that is, is the only way we

21 have in terms of making our voices heard. But are

22 there any other suggestions? The advisory board may

23 not be it but are there any other ways you suggest

24 in terms of the public having a say?

25

2   JONATHAN AUERBACH KAEHNY: I'll tell you

3   what I think is by far the most important thing to

4   make this law work. Two things; one, agencies need

5   to understand, Mike Flowers alluded to this, that

6   this law helps them. It reduces their work load. It

7   reduces public complaints, FOIL requests,

8   irritation. And secondly city council needs to use

9   this law. City council staff need to use this law

10   whenever city council is asking for information

11   from an agency which is every single day in

12   oversight hearings they need to say can we get this

13   out of a, a data set that is in the open data

14   portal instead of having to go through an agency

15   and in ad hock request way. And when you guys ask

16   for it and when you are, and your central staff

17   starts using it all the time and taking it for

18   granted and council members in every single

19   committee starts saying is that in the open data

20   portal. Then this law will work. I think those are

21   the two most important things. And in, in terms of

22   council maybe there's a discussion with central

23   staff that can happen about the open data law and

24   about the opportunity it presents. Because part of

25   this law is empowering the legislative branch with

2   information because as we know dragging information

3   out of the executive, out of the mayor is a big

4   challenge for you as well as the public. So this

5   law is for you and you have to use it. And I would

6   say that's, you know if we can make you do that

7   that is more powerful than any advisory panel or

8   anything else.

9   COUNCIL MEMBER BREWER: Excellent. Next

10   question is, this is… Oh go ahead sorry.

11   RACHAEL FAUSS: I, I was just going to

12   add, I mean we were talking, you mentioned earlier

13   about not all the incoming members of the council

14   knowing about the open data law. I think having the

15   oversight hearing early in the next year right

16   after the next compliance plan comes out I think

17   shows… [interpose]

18   COUNCIL MEMBER BREWER: Council member

19   Cabrera can make sure they all show up.

20   [laughter]

21   RACHAEL FAUSS: Yeah but I think it's an

22   opportunity to educate the new members about the

23   law and its, its benefits.

24   GENE RUSSIANOFF: Is this included in

25   any of the training for the new members? I don't

2   know it was a late suggestion but you know if

3   they're, they should be familiar with the tool

4   their arsenal for doing oversight of the mayoral

5   agencies?

6   CHAIRPERSON CABRERA: Actually it's not

7   because you know it wasn't a, it wasn't law but I,

8   the new members, they haven't been briefed yet.

9   They haven't been trained. And, and, and the

10 various pieces that I went through four years ago

11 but it's a good idea and it will definitely pass on

12 to our, the next speaker.

13   NOEL HIDALGO: I just want to point out

14 that Pediacity is going to be testifying here as a

15 BigApps winner. As a current participant of

16 improving the New York City open data law they have

17 a unique tool that actually provides that type of

18 insight. And that's something that's coming from

19 the citizens right now. If DoITT can't do it we're

20 more than happy to do it as citizens. So just give

21 us the data and we'll build those dashboards for

22 you.

23   COUNCIL MEMBER BREWER: Final question

24 is, this is a little bit broader but, one of my

25 concerns is making sure communities be they parent

2   organizations, community boards, not only have

3   access want to know how to work with the data.

4   Otherwise it's like the city council. We have to

5   use it. So my question would be do you have some

6   suggestions as to how the communities in the

7   broadest sense could be taught, shared with,

8   supported in terms of using that data. Obviously it

9   has to be in some cases you know mapped and shown

10 how to do overlays and how it's relevant to the

11 communities, etcetera. And I, I, you know this is

12 obviously something that I care deeply about but I

13 also think that's how the portal gets used and gets

14 improved.

15   NOEL HIDALGO: I'll take that again. So

16 that's one of the things that we do through Code

17 for America's brigade program. We do it here

18 locally through Beta NYC and we do it nationally,

19 actually internationally through 30 different, 32

20 different community groups. So here in New York

21 City we meet on a weekly basis and we're more or

22 less getting the technologists who can understand

23 this data to build platforms like Pediacity does so

24 that way you can get that insight. And then the

25 next step is to figure out that curriculum around

2   education. We hope at the, with this unique

3   opportunity with a new council coming in that we

4   can actually be going to all the council members

5   not only showing that we are constituents but that

6   we're also building these tools that provide that

7   type of insight. Other great opportunities would be

8   to bake it into the curriculum within education.

9   Education, public education either through library

10 centers, educate, like actual formal education

11 program, having 21st century civics programs making

12 kids required learning on how to crunch the data,

13 have them understand their particular community. In

14 boy scouts I did that. You know why can't I do that

15 now have we no computer you know? So those are

16 steps that the next council needs to take.

17   COUNCIL MEMBER BREWER: Thank you.

18   JONATHAN AUERBACH KAEHNY: An incoming

19 borough president could do something cool here.

20   [laughter]

21   COUNCIL MEMBER BREWER: Yeah income

22 borough president is thinking about it but I'm

23 getting your ideas.

24   JONATHAN AUERBACH KAEHNY: Okay which is

25 to get the community board district managers and

2   committees to identify what their most requested

3   data is and then to collate that into actual data

4   sets.

5   COUNCIL MEMBER BREWER: I'm hoping Code

6   for America will help us do that.

7   JONATHAN AUERBACH KAEHNY: They're going

8   to help you.

9   COUNCIL MEMBER BREWER: Thank you.

10   GENE RUSSIANOFF: I just wanted to

11 confess that I have very little ability to deal

12 with the data. I can find it but then knowing how

13 to play with it is very difficult so various

14 educational opportunities would be greatly

15 appreciated.

16   COUNCIL MEMBER BREWER: Me and Gene need

17 help. Thank you.

18   [laughter]

19   CHAIRPERSON CABRERA: Thank you so much.

20   GENE RUSSIANOFF: Thank you.

21   CHAIRPERSON CABRERA: We're going to

22 have the next panel. Ellen McDermott from

23 OpenPlans, Dominic Marro NY Tech Meetup, Nathan

24 Storey from Ontodia, and Juan Martinez from

25 Transportation Alternative.

2   [pause]

3   CHAIRPERSON CABRERA: Hi, you may begin.

4   ELLEN MCDERMOTT: Well it's in front of

5   me so I have to begin.

6   [laughter]

7   ELLEN MCDERMOTT: Good afternoon. My

8   name is Ellen McDermott. I'm co director of

9   OpenPlans a not for profit technology development

10   firm based in Manhattan. We work with city and

11   state agencies, community boards and community

12   groups to provide them with better tools for

13   transportation planning and neighborhood planning.

14   OpenPlans worked closely with council member Gale

15   Brewer and others on the council to get the open

16   data law passed. We've been a member of the New

17   York City Transparency Working Group since its

18   inception. In my testimony today I will give

19   examples of how open data works, supports our work

20   as a local technology developer but I also have

21   some recommendations to strengthen the open data

22   ecosystem in New York City. This fall we have been

23   working with community board six in Brooklyn and

24   CB6 and CB7 in Manhattan to create maps showing

25   requested capital spending. With council member

2   Lander's office and the participatory budgeting

3   project we are helping residents of eight council

4   districts be more involved in local spending

5   decisions. Just last week we created a map for the

6   community group Make Brooklynn Safer which they are

7   using to gather information on street safety

8   issues. The data they are gathering is informing

9   conversations with the 88th precinct as they work

10   together to improve street safety after several

11   recent pedestrian fatalities. The examples I gave

12   are all made easier with the city's open data. We

13   would like to see more data to enable more

14   community based planning. I have several

15   recommendations. First, release more data about

16   spending and capital improvements. Community

17   organizations will be able to advocate more

18   effectively with access to more detailed, up to

19   date, information about capital spending data.

20   Specifically agency responses to community board

21   requests past and future discretionary spending by

22   council members and information about planned and

23   proposed capital spending by the DOT, the Parks

24   Department, and other agencies making physical

25   changes to communities.

2   Secondly, work with users to improve

3   the open data portal. DoITT should involve users of

4   the data portal in design, design decision, design

5   decisions to ensure that the portal is working for

6   its intended audience. By convening a usability

7   clinic Do, DoITT can tap into the wealth of

8   experience available in New York's policy and

9   software communities to guide decisions and make

10   the portal even more useful. This effort needs to

11   be regular and ongoing. Third improve city data

12   through community collaboration. The process of

13   opening up data so far has been one way. Data from

14   city agencies has been published in formats that

15   people can use. Looking into the future we

16   encourage this committee to consider how agencies

17   can save money and have up to date information by

18   using community edited data. For example community

19   mapping of street trees and the freely editable

20   open street map. The recent collaboration between

21   DoITT and volunteer cartographers to include

22   building footprint data into open street map is a

23   commendable example. Implementing two way sharing

24   of open data raises complex issues about data

25   quality and data management. And I urge this

2   committee to charge DoITT with investigating this

3   topic further.

4   Finally I welcome this hearing and

5   commend the committee for giving us and other data

6   users a forum to share our concerns and ideas. I

7   suggest regular oversight hearings on progress by

8   DoITT and agencies in opening up their valuable

9   data. My coworkers and I are always available to

10   assist committee members with this important topic

11   to support greater community involvement in making

12   New York City a better place to live and work.

13   Thank you.

14   NATHAN STOREY: Good afternoon honorable

15   council member. I am Nathan Storey, Product Manager

16   for Pediacities[dot]com encyclopedia of city data

17   made by Ontodia, a local made in New York company

18   and winner of your previous BigApps. I'm also a

19   member of Beta NYC, Code for America brigade in New

20   York City and I strongly support the city's open

21   data efforts and acknowledge that the city has made

22   great progress thus far as other people have

23   testified. But I know we must do more to realize

24   the full economic potential of open data. In a

25   study released just last month the McKinsey Global

2   Institute estimates, estimates that open date has

3   the potential to unlock three to five trillion

4   dollars in economic value per year in the US alone.

5   In New York much focus has been placed on using

6   data contests known as hackathons to unlock that

7   potential. Hackathons are intended to spur the

8   creation of projects that turn into viable

9   businesses. And while that is certainly the case

10 with Pediacities our company and my job are the

11 exception. To create more open data jobs we need

12 stronger ongoing partnerships between the producers

13 and the users of city data. First we need agencies

14 to replace their closed internal data systems with

15 open data platforms. This would create open data

16 jobs of course but it would also create better

17 feedback loops to correct errors, allow for more

18 efficient identification of gaps in the available

19 data sets, and make possible the incorporation of

20 community generated data. Second we need to invest

21 more resources in the expansion of open data across

22 the civic realm. More integration between DoITT's

23 open data team and community groups and an

24 expansion of volunteer Code Corp projects beyond

25 the disaster response are good places to start. In

2   addition to contests we should invest governmental

3   and philanthropic resources in imbedding civic

4   technology experts directly in our city's community

5   boards, city council staffs, and agency rosters.

6   Finally the city council and new administration

7   should look for ways to create and track ambitious

8   performance indicators such as early indicators for

9   risks of foreclosure, gentrification,

10   disinvestment, climate vulnerability, and the like.

11   If open data can support the creation of these

12   kinds of early warning systems the economic and

13   social impact will be profound. Thank you.

14   DOMINIC MARRO: Good afternoon. My name

15   is Dominic Marro Marrow. I'm a Staff Attorney at

16   Reinvent Albany and I'm here today to read a brief

17   statement on behalf of the New York Tech Meetup

18   which was unable to send a representative to

19   today's hearing but wanted to emphasize its strong

20   support for the New York City open data law. The

21   New York Tech Meetup is a 36 thousand member

22   nonprofit organization that supports the growing

23   and wide technology community. Our goal is to help

24   build a sustainable technology industry that drives

25   economic growth, leads innovation, and creates

2   positive high impact change for our local

3   communities and the world. The New York Tech Meetup

4   has been integral to the birth of New York's new

5   technology industry with nearly every notable new

6   company making its debut at the New York Tech

7   Meetup's monthly event. Companies that have demoed

8   at the New York Tech Meetup in the past include

9   Foursquare, Livestream, Tumbr, Vimeo, Aviary,

10   Audible, Twilio, Aereo, Skillshare, Shapeways,

11   TheLadders, HowAboutWe, Codecademy, and many many

12   many others. The tech meetup strongly supports the

13   New York city open data law. The law is a powerful

14   tool for opening up government, improving public

15   services, and promoting both economic development

16   as well as tech jobs. A key feature of the law is

17   that it endorses the basic values expressed in the

18   Freedom of Information Law, that government

19   information, in this case electronic information or

20   data, belongs to the public. Our government curates

21   that information but we the public own it. The law

22   recognizes that the public you, me, and people like

23   our members can take that data and find useful and

24   productive things to do with it. Like smartphones

25   open data is part of a new era in which technology

2   tools are changing how we live and how our

3   government works. We think the open data law is

4   working and has huge positive potential for New

5   York City. We look forward to working with council

6   and the next mayor to unleash that potential for

7   everyone's benefit. Thank you.

8   JUAN MARTINEZ: Hi there. Thank you to

9   the chair and the committee for the opportunity to

10   testify. Juan Martinez with Transportation

11   Alternatives. We're a nonprofit, been around 40

12   years, we work to make the city better for walking,

13   biking, public transit. And I wanted to talk to you

14   from the perspective as a, as a consumer of the

15   data that you've unlocked and how important it is

16   for us to make sure that it's done well. The jest

17   of it is that local law 11 of 2011, a bill authored

18   by council member Lappin required the NYPD to

19   publish crash data and moving violations data. It's

20   a huge opportunity to encourage the agencies to

21   more efficiently focus their street safety

22   resources and to empower communities to make

23   decisions about their own streets. The, the problem

24   is that the promise of the bill has been frustrated

25   right? We anticipated that council members,

2   community board members, community members,

3   community advocated would all be able to read this

4   data and, and see what was going on, identify

5   danger patterns and trends. But the way that the

6   data's been published by the NYPD has frustrated

7   that goal. And that's a shame because the demand

8   for the data remains very high. The, the need is

9   still very high. And if it weren't for the efforts

10 of a few civic hackers, highly talented, very

11 diligent, hardworking, private citizens who did

12 what they could to unlock the data from the way

13 that NYPD published it we wouldn't see any benefit

14 from the law that council member Lappin passed. But

15 because of their efforts we have been able to get

16 some of the data out. Right? It's not perfect. It's

17 not as clean as it could be but it's usable. And as

18 a result we're able to work with journalists, work

19 with council members who are curious about a

20 particular intersection, and that kind of thing.

21 But it's very time intensive, very labor intensive,

22 and people are coming to us for the information as

23 opposed to being able to find it themselves. Your

24 staff should be able to online and find this data.

25 And I've, you know receive calls from council and

2 staff all the time. Happy to do it. Happy to do it

3 but it's not the ideal way to do it. It's not what

4 was anticipated. Now the thing about it is that

5 it's not as if we need the NYPD to do more work in

6 order to make local law 11 what it could be, make

7 this publication of the data what it could be.

8 Actually we're asking them to do less work. What

9 they do is they publish the data in a format. They

10 add formatting to it in a way that introduces the

11 errors and makes it so that somebody can't load it

12 into their computer and produce a map with a, with

13 a, with a push of a button right. They add to the

14 data. They, they, they, they obfuscate it, they

15 confuse it right. So what we hope is one of the

16 outcomes of this hearing and council member Lappin

17 has introduced a bill modeled on council member

18 Cabrera's crime mapping bill. What we're hoping at

19 the end of this is you can convince the NYPD to

20 simply do what they're doing but put a little less

21 effort into it and make it something that we can

22 use. Well and something that this community can

23 use. Because the really, the potential's immense.

24 The idea that you could have a report sent directly

25 to your email telling you where the hotspots in

2 your neighborhood are. The next time you talk to

3 DOT and you're saying I need a speed bump you can

4 tell them why you need a speed bump because you

5 know how many crashes happened at that

6 intersection. Your community board members would

7 stop, would be able to focus on the most important

8 intersections as opposed to the ones where, have

9 gotten the most attention lately. It's a huge, huge

10 benefit if they would just do a little less work.

11   CHAIRPERSON CABRERA: Thank you so much.

12 I was going to ask you. I keep hearing the NYPD. Do

13 you think, do you think this is beca… Is there a,

14 do you see an intention on the part of

15 administration to present the information this way?

16 Or, or is because they don't know any better? Or is

17 it because they're getting ready to roll out the

18 crime map bill and they don't want to take you know

19 away the big day when it comes out. Because with

20 that information right now you can actual

21 divel[phonetic] up your own right? If you had, so,

22 what do you, what do you see? How do you interpret

23 this?

24   JUAN MARTINEZ: There are very

25 sophisticated agency. They are a very big agency, a

2 four billion dollar agency. They have the staff to

3 do it well and there's been an intense scrutiny

4 from council member Lappin, council member Vacca on

5 this particular point. I hope that what they want…

6 There was actually a hearing on council member

7 Lappin's bill a few weeks ago. And the

8 administration's position was that people in the

9 public can map the data today. And that is true but

10 because of the errors that are introduced by their

11 formatting it's not as good as it could be. It's

12 far from as good as it could be. So I hope that

13 they take the opportunity to really hear what

14 council member Lappin and other members of the

15 councilor saying which is that you want this data

16 to be available in its purest form. That's one of

17 the things. You know they're, they're very very

18 smart people who know their stuff very well. But at

19 the end of the day what I take away from a lot of

20 it is that they just want the purest form of the

21 data before it gets mucked up by well-intentioned

22 people.

23 CHAIRPERSON CABRERA: Council member

24 Brewer.

25

2   COUNCIL MEMBER BREWER: [off mic] I'm

3   just wondering…

4   COUNCIL MEMBER BREWER: Do you make

5   comments on the data, regarding the data, or do you

6   know authors that do? Do you think this is

7   effective? Do you get answers? Because obviously

8   we're trying to improve it and the one way we can

9   do that is by the public. We were not able to have

10 any other mechanism. So I'm just wondering if

11 you've looked at that if you, you know and if there

12 is some improvement to that process.

13   JUAN MARTINEZ: So DoITT has done a

14 great job at being at ad events where civic hackers

15 are and being open to talk about the data sets and

16 they have made these tools available for commenting

17 on individual data sets. You know sometimes I think

18 it's a resource issue of like you know you have to

19 prioritize what you're doing so they don't always

20 get around to the individual comments. But they do

21 get around to a lot of them. So I commend them

22 about that. I think there's a, there's a deeper

23 problem which was alluded to you about getting

24 community input to fix individual kind of entries

25 in a data set and that introduces a lot of data

2   quality issues and, and, you know the process for

3   that is a lot more complex. But I, I agree that we

4   should, the council should task DoITT with looking

5   further into how we can get even better input to

6   get communities to improve data sets about

7   themselves.

8   COUNCIL MEMBER BREWER: Thank you very

9   much.

10   CHAIRPERSON CABRERA: Thank you so much.

11 Appreciate your time and your information. Calling

12 upon the last panel; Sara LaPlante from New York

13 City's Liberties Union, Lourdes Cintron from City

14 Wide Mental Health Project, Matthew Bishop from

15 IGiveMore[dot]org. Hmm, I like that name. Patti

16 Banghart Children's Defense Fund.

17   [pause]

18   CHAIRPERSON CABRERA: Whenever you're

19 ready you may begin. Whoever feels brave.

20   LOURDES CINTRON: [off mic]I'm the one

21 with the least experience here in technology so

22 maybe I should be the one to go first and get it,

23 get it over with.

24

25

COMMITTEE ON TECHNOLOGY   93

2   CHAIRPERSON CABRERA: Thank you. If you

3   could just get the microphone closer. Appreciate

4   that.

5   LOURDES CINTRON: Alright.

6   CHAIRPERSON CABREAR: Go ahead.

7   LOURDES CINTRON: My name is… Good

8   afternoon my name is Lourdes Cintron, the founder

9   of the City Wide Mental Health Project a graduals

10 group of recipients of mental health services,

11 friends, and relatives. And we are striving to

12 empower consumers to learn to money toward the

13 quality of services they receive in their community

14 mental health's programs among other goals. We

15 appreciate this opportunity you grant us to

16 contribute our feedback on the topic of the, the,

17 New York City open data initiative. There is no

18 data on mental health issue in that website. I went

19 there trying to test the website to see how it

20 works. Technical aspect that I notice is that the

21 website is confusing and in my view as it is now

22 useless for the purpose stated in the law. For

23 example the icons for selecting the data format, a

24 format presentation are not user friendly, meaning

25 they are not enjoyable. It requires high levels of

2   computers and research skills to figure out which

3   one to select and once selected the format is still

4   confusing. I could not use it even though I do have

5   a little bit of computer skills. As it is now most

6   of the members in my group do not have the skill to

7   navigate the website graphical user interface so

8   called. If they needed to access the information

9   supposedly at they level. This website was designed

10   for researchers not for the public in general.

11   Transparency issues; searching in the data

12   catalogue box for mental health data is an exercise

13   in futility. The search for either mental health or

14   Department of Health and Mental Hygiene give you

15   both of them quote unquote use your famous babies

16   names. And food vendors without permits. Pages and

17   pages of that kind of information repeated. Also a

18   search on 311 reports shows not a single call

19   requesting information about mental health services

20   or a single incident related to it. Almost all 311

21   reports seems to tell something. I'd relate it to

22   vermin and rats. A researcher could easily conclude

23   that ratting firm infestation has no impact in the

24   city's mental health. This could matter for policy

25   and budget purposes. Finally how is transparency

achieved. By publishing the name and addresses who

commit a minor infraction. For example names of

street vendor without permit, names and address of

people without a vending industry without permit.

But neglecting to publish the name of big city

company CEOs who have the defrauded the city in

millions of dollar like Mark Masors[phonetic] of

City Times[phonetic] whose name is not there for

example. If transparency is going to be depended on

report that meaningless data and if quoting from

the law the city lost not warrant the completeness,

accuracy, content, or fitness for any particular

purpose of the data publish. Then this new law is

nothing but a let me just call it a mistake. Let's

be realistic. There has been not a single

government agency in this nation history which

welcomes transparency especially since 9/11. This

website as it is now is the latest edition to that

historical fact. And I would like to mention two

more things that came to mind. People talk about

all kind of that data here nobody mentions the

Bronx, we don't exist, we are not in the map. And

the other thing is that that data is not

2   transparency to meet transparency is a point of

3   view, a desire to do the right thing. Thank you.

4   [pause]

5   MATT BISHOP: Technology issues.

6   [laughter]

7   MATT BISHOP: My name is Matt Bishop.

8   I'm founder and CEO of IGiveMore[dot]org. We're a

9   crowd funding platform for nonprofit organizations.

10 I'm also a member of the board of directors of

11 Citizen Action of New York and I previously worked

12 at volunteers of America which had about a hundred

13 million dollars in contracts with 25 different

14 government agencies so I've worked very closely

15 with government agencies in terms of understanding

16 how nonprofits are working with government agencies

17 to deliver services to the homeless and the poor.

18 I, I want to thank the city council and the

19 Department of Information Technology for their

20 leadership on this issue. It is clear that New York

21 City is a national leader on this issue for open

22 data and civic hacking. I think the largest

23 unfulfilled opportunity of New York City open data

24 is not just the ability to, for journalists or

25 agencies to or individuals to build apps or make

2   that governments data more accessible. The

3   opportunity really is for agencies to create

4   laboratories of policy analysis by which third

5   party apps can collect information and apps can be

6   built to assist agencies and policy analysis and

7   program management. I mean the phones that we have

8   in our pocket are more powerful than the cell phone

9   that you know our president used 10, 12 years ago.

10   So I think I'll, I'll just give a few really quick

11   examples of ways that, that, that could work. So

12   there are a number of devices and mobile apps that

13   provide APIs which could push health care data to

14   some kind of government agency for policy analysis.

15   For example Fitbit or 1Keeper. You know these,

16   these apps can track calories burned, heart rate,

17   and other health data and this could be used to

18   improve health outcomes for citizens. It can form

19   medical treatment by doctors by syncing with e-

20   health records like the common health portal that

21   is being created by the, the state Department of

22   Health. This, and this could be used with a, so, so

23   what would that look like. So city agencies could

24   create policy challenges by which they're trying to

25   collect certain data and then civic, and gage

2   developers and hackathons to, to basically leverage

3   the power of mobile applications to, to, to collect

4   that data. And then also APIs could be allowed to

5   push data to, to city agencies to help streamline

6   the process of applying for government programs. So

7   for example imagine you're about to submit your

8   taxes online, like on TurboTaxed or something like

9   that. And above the submit button is something that

10 says would, would you like to see what government

11 programs you qualify for? Click here to see what

12 government programs you qualify. You click submit

13 and then a page opens that lists you know you

14 qualify for food stamps, you qualify for NYCHA

15 Housing whatever it might be. Would you like to

16 submit your information to these agencies to apply

17 for these services? You know they could very

18 quickly review what government programs they

19 qualify for and then apply to them directly. While

20 that may not be the end of the application process

21 that could be the beginning process of submitting

22 that information to the agency before paying a

23 visit to HRA or NYCHA or whatever else is involved

24 in the application process. This could completely

25 streamline the, the way in which, and, and citizens

2   are engaging with government and streamline the,

3   reduce the amount of paperwork for these agencies.

4   And then finally I think creating an API that

5   allows for public apps to collect data provides an

6   opportunity for collaboration amongst government

7   agencies not just between agencies but between

8   levels of government. For example the white house's

9   innovation fellows created, they have created a,

10 open source apps that city agencies around the

11 country can use for common authentication. Right?

12 Because part of the problem of open data and, and

13 agency collaboration is that they all operate on

14 different databases and systems. Right? So what the

15 white house is doing to address this is to say

16 here's a common authentication system so that

17 government data multiple agencies across government

18 levels could provide, could use common

19 authentication systems so that consumer needs are

20 met around who they are as opposed to you know I

21 have one health, one portal by which I log in for.

22 You know New York City Department of Health, one

23 portal for New York state Department of Health, and

24 one for the federal government. Right? So these

25 are, there's a, there's a real opportunity in

2   collaborating with different agencies on different

3   government levels to make government relevant to

4   you know and more, more accessible to your everyday

5   citizens. So I think these are just a few of the

6   applications that could be realized by having data,

7   having the next level of APIs pushing data to

8   government agencies not just making their data

9   accessible.

10   [pause]

11   PATTI BANGHART: Hi I'm Patti Banghart

12 and I'm the Director of Early Childhood Policy at

13 the Children's Defense Fund of New York. I guess

14 this afternoon I'll tell you a little bit about our

15 advocacy efforts and how we think the open data law

16 would help us be more effective advocates. The

17 Children's Defense Fund Leave No Child Behind

18 mission is to ensure that every child has a healthy

19 start, a head start, a fair start, a safe start,

20 and a moral start in life, and successful passage

21 to adulthood with the help of caring families and

22 communities. As part of our advocacy efforts we

23 launched the CDF Cradle to Prison Pipeline Campaign

24 which is a national call to action to stop the

25 funneling of thousands of children, especially poor

2   children and children of color down life paths that

3   often lead arrest, conviction, incarceration, and

4   even death. The Children's Defense Fund of New York

5   office currently engages in policy analysis and

6   advocacy with regard to early childhood education,

7   healthcare, education, justice, and youth justice.

8   We work closely with a number of community based

9   organizations as well as faith leaders, youth

10 groups, and parent groups to ensure that our

11 advocacy is shaped by the everyday realities of our

12 most vulnerable New Yorkers. In terms of the New

13 York City open data law and how it can help further

14 social change and progress for our city's children

15 and families I'm just going to outline a couple of

16 ways that the law would further help with policy

17 analysis and advocacy. With respect to early

18 childhood education and programming services in New

19 York City the administration for children services

20 currently doesn't publicize data on the child care

21 of subsidy and early learn programs. Therefore

22 little is known about who is being served the type

23 of care that is chosen and the quality of service

24 being provided. Having public data that can speak

25 to these areas, I admit it would be extremely

2 helpful for determining program access and program

3 quality. For instance while the early learn program

4 posts data on overall program enrolment there is no

5 breakdown to identify the demographics of enrollees

6 the specific catchment areas, where enrolment is

7 happening. Additionally any classroom observation

8 data that is being collected is not publicized as

9 well. As child advocates it is our hope that this

10 data and other information on program related

11 outcomes that will be collected in the future could

12 be made public. Making data collected on early care

13 and education services more transparent and

14 accountable to achieving specific measures for the

15 wellbeing of children in the earliest years of

16 development is a necessary step to ensuring that

17 all children have a head start in life. With regard

18 to the tracking and monitoring of disciplines used

19 in New York City schools by the Department of Ed.

20 And the New York Police Department, while the

21 student safety act requires data is released on

22 school safety and disciplinary issues there is a

23 significant portion of data that is redacted and

24 incomplete including incidents involving arrests

25 and suspensions of students. While the data

2   currently provided has allowed child advocates to

3   shed light on how students of color are

4   disproportionately being fed into the school to

5   prison pipeline there is more that needs to be

6   revealed. Both the DOE and the NYPD need to be held

7   accountable for the, by the community to ensure

8   that all data released is complete and accurate.

9   Concerning measuring the public health of New York

10   City's children and family data detailing health

11   outcomes of city residents is highly important in

12   determining the need for resources and the

13   effectiveness of specific interventions. For

14   example openly providing data that measures health

15   outcomes for children in youth who receive health

16   services in schools by school nurses or school

17   based health center providers allows for

18   determinations to be made about treatment and

19   program advocacy when it comes to measuring impact

20   and outcomes. With regard to youth justice data

21   currently collected on children placed in secure

22   facilities have clearly demonstrated that children

23   are regularly restrained, injured, and confined to

24   their rooms in these settings. This is a

25   significant finding that requires action. Access to

2   data that can show such a trend is key to

3   demonstrating the need for system reform. If we are

4   to have any hope of reducing these harms we must

5   not allow such harmful practices to proliferate in

6   the dark hidden from the public. Implementation of

7   the open data law supports this transparency

8   allowing for accountability and reform.

9   Finally in closing I just want to point

10 to an example of how the intent of the open data

11 law has been recognized as necessary by the New

12 York City council. In 2013 new legislation was

13 passed regarding reporting requirements for ACS

14 with regard to youth detention and placement. This

15 new legislation replaced previous legislation

16 requiring ACS to report on demographic and incident

17 related data in secure and non-secure detention

18 facilities housing children. The new legislation

19 added in the inclusion of non-secure placement and

20 limited secure placement facilities. Demographic

21 data collected now includes all four categories

22 including secured detention, non-secure placement,

23 and limited secure placement. Additionally for most

24 categories it encompasses a comprehensive spectrum

25 including age, gender, race, zip code, and charge

2   data. Expansion of data collection to include all

3   juvenile facilities under ACS jurisdiction is a

4   necessary step to monitoring systems and

5   determining how to provide optimal intervention to

6   youth who are being detained and placed. The intent

7   of the open data law calls for such

8   comprehensiveness and ultimately comprehensive data

9   is essential to providing evidence on how to serve

10 the profound needs of youth who enter the justice

11 system. Thank you very much.

12   [pause]

13   SARA LAPLANTE: Hi my name is Sara

14 LaPlante and I'm the Data Analyst at the New York

15 Civil Liberties Union where I gather and analyze

16 data to support our program initiatives. In

17 advocating for the rights of New Yorkers were often

18 rely on publically available data and records

19 provided under the freedom of information law.

20 Access to information such as the NYPD's Stop and

21 Frisk Database has led to major policy changes. For

22 example in the most recent reporting period 80

23 thousand fewer individuals were stopped and frisked

24 by the NYPD. Public debate about these practices

25 was fueled by data demonstrating their

2   ineffectiveness and unbalanced racial impact. The

3   NYCLU wholeheartedly supports the open data law but

4   we would like to see DoITT improve implementation

5   and effectiveness. DoITT has failed to compel city

6   agencies' full compliance in reporting data and

7   mandated by the law. This is particularly evident

8   with city agencies known to shirk compliance with

9   open records laws such as New York Police

10 Department. In September DoITT released its open

11 data plan that included a list of all data sets to

12 be published pursuant to the law. The NYPD an

13 agency with the fifth largest budget in the city

14 and perhaps the most public attention of all

15 government offices submitted the names of only six

16 data sources that it deemed public as defined by

17 the law. These six data sources are image and text

18 heavy reports published by the NYPD. While they

19 contain important data summaries and graphs and

20 tables submitting entire reports to the web portal

21 misses the mark of the law. PDF files are not

22 easily machine readable and the use of the embedded

23 data tables require intensive labor from anyone

24 hoping to analyze the data in a way different than

25 presented in the department's reports. Further

2   these reports come prepackaged with a public

3   relations angle from the NYPD. Researchers using

4   them are left to deconstruct messages from the

5   provided narratives and reconstruct their own

6   conclusions. Finally even if the reports the NYPD

7   plans to provide under the open data law were raw

8   data provided in a machine readable format the list

9   they proffer is far from exhaustive. For example

10 the NYPD already publishes its stop, question, and

11 frisk database, a database reporting information on

12 each terry stop made by a NYPD officer. However

13 it's in a format inaccessible to the public yet the

14 department excludes this database from the data

15 sets it plans to publish on the web portal. There

16 are a number of other NYPD data sets excluded from

17 the open data plan that can and should be published

18 including; the locations of crime incidents and

19 reports, information on summonses issued for

20 noncriminal violations, and operations of the

21 school safety division. We include a list of these

22 data sets, things that we would like to see, in our

23 written testimony. The NYCLU has the following

24 recommendations for the council today. The city

25 council must continue to fund local law 11. Without

2   funding DoITT will not be able to maintain the web

3   portal established by the law and without the web

4   portal the law loses its best tool for sharing data

5   from city agencies. Interpret the law broadly. To

6   ensure the transparency and accountability of the

7   heart of the law city agencies should report any

8   public data set they maintain. The default should

9   be to include more data rather than reporting only

10   the minimum data sets required by a narrow

11   interpretation of the law. City agencies that

12   comply fully and broadly with the law will save

13   resources used in responding to FOIL requests by

14   proactively promoting transparency. And finally

15   DoITT and the city council must hold city agencies

16   accountable for the data sets they plan to report.

17   DoITT should review more closely the data sets in

18   the open data plan and reach out to underreporting

19   agencies to encourage a broad inclusion of public

20   data sets. This may require additional follow-up

21   with agencies that are known to obfuscate open

22   records requirements such as the NYPD and also the

23   DOE. DoITT and the council should conduct

24   compliance audits and work with the new

25

2   administration to usher in a new era of

3   transparency for New York City.

4   CHAIRPERSON CABRERA: Thank you so much.

5   I, I had a question. You mentioned something about

6   the Bronx. I'm from the Bronx so you don't have to

7   ask. You, you mentioned that there is no data…

8   [interpose]

9   LOURDES CINTRON: No, no what I…

10 [interpose]

11   CHAIRPERSON CABRERA: …from the Bronx.

12 Did I misunderstand you.

13   LOURDES CINTRON: Yes, I wasn't too

14 clear.

15   CHAIRPERSON CABRERA: Okay.

16   LOURDES CINTRON: Because I realize that

17 as… [interpose]

18   CHAIRPERSON CABRERA: If you could just

19 bring the microphone closer to you.

20   LOURDES CINTRON: I realize that as I

21 was sitting here listening to the other people

22 talking and doing presentations that all of the

23 maps that they were showing about… [interpose]

24   CHAIRPERSON CABRERA: Oh I got you.

25

2   LOURDES CINTRON: …data, it was all…

3 [interpose]

4   CHAIRPERSON CABRERA: …it's typical.

5   LOURDES CINTRON: Manhattan… [interpose]

6   CHAIRPERSON CABRERA: Yeah it's typical.

7   LOURDES CINTRON: …and yeah.

8   CHAIRPERSON CABRERA: I, we're…

9 [interpose]

10   LOURDES CINTRON: The Bronx… [interpose]

11   CHAIRPERSON CABRERA: …we're are not

12 used to it either.

13   LOURDES CINTRON: You know it's like we

14 don't exist you know.

15   CHAIRPERSON CABRERA: Yeah, I, I hear

16 you. I always complain about that.

17   LOURDES CINTRON: But anyway we're the

18 poorest borough in the nation so that's…

19 [interpose]

20   CHAIRPERSON CABRERA: Yeah poorest

21 congressional district in the nation.

22   LOURDES CINTRON: Yep.

23   CHAIRPERSON CABRERA: We, we need to

24 give it all the attention that we need. Okay. Great

25 great I just wanted to… And then you're, you're

2   saying that there is no, I'm a licensed man to help

3   councilor and my doctorate's in counseling so this

4   is an area that means a lot to me. You're saying

5   that there is absolutely no data on mental health

6   issues that is viable to the field, that is could

7   be useful for the field.

8   LOURDES CINTRON: There is none, period.

9   [crosstalk] It's not that it's not useless, that is

10 none.

11   CHAIRPERSON CABRERA: None.

12   LOURDES CINTRON: If you know if it were

13 something useless I would have said well that's

14 something but it's useless.

15   CHAIRPERSON CABRERA: Okay.

16   [laughter]

17   LOURDES CINTRON: To the program that is

18 nothing.

19   [laughter]

20   LOURDES CINTRON: There is nothing.

21   CHAIRPERSON CABRERA: There is nothing.

22   [laughter]

23   CHAIRPERSON CABRERA: Just rats?

24   LOURDES CINTRON: Just rats.

25   CHAIRPERSON CABRERA: Rats.

2   LOURDES CINTRON: Pages and pages.

3   CHAIRPERSON CABRERA: Okay.

4   LOURDES CINTRON: And, and, and it's

5   important to us obviously. The reason what I

6   mentioned our goals in our, in a grassroot group is

7   that we are one of those group that is trying to

8   denounce whether we can see there is a culture of

9   abusing the mental health system. And we need to

10   document what we are talking about. It's not only

11   the histories that we can bring personally. The

12   painful histories which people don't pay attention

13   to us because we are stigmatized. Whatever we have

14   to say is useless.

15   CHAIRPERSON CABRERA: Right. So you the

16   data… [interpose]

17   [crosstalk]

18   LOURDES CINTRON: Because we are crazy

19   people.

20   [crosstalk]

21   LOURDES CINTRON: So I need to look for

22   this and then they come and present like usually oh

23   this is beautiful come here and you go there and

24   that's nothing. The department of mental health

25   citywide or statewide they are one of the, the, the

2 agencies that like to not tell you what is really

3 going in our programs. And that's why you find now

4 you go on taping, typing they are department of

5 health you find nothing, nothing, nothing.

6   CHAIRPERSON CABRERA: Okay. Thank you so

7 much. Council member Brewer.

8   COUNCIL MEMBER BREWER: I just want to

9 follow-up on that because I have a big push for

10 years getting more culturally appropriate mental

11 health in the schools and, and that list is hard to

12 come by. So it, not only is it not on the database

13 I was just checking but also it's not available

14 period. So I think, I mean it exists we just can't

15 get it. So I think we will all work on the mental

16 health challenges in terms of some of the

17 databases. What would be helpful though would be

18 either to the chair or to the comment periods in

19 the data to indicate what kind of mental health

20 data bases you might like as opposed to just you

21 know. I know I want the one from the schools as an

22 example but there may be others that you could find

23 helpful and I think if you shared that with the

24 chair that would be helpful. Okay. My other quick

25 question for generally and then that's the only one

2   I have is what do you… A lot of the databases are

3   there but not in a format that is helpful. So one

4   of my questions is it is supposed to be in a format

5   that is helpful just so you know but how much do

6   you find that to be a challenge number one. Two,

7   have you made comments that have or have not been

8   responded to. And just three what suggestions do

9   you have because you're often in the communities to

10 help the community boards, the parent associations

11 etcetera use the data? What would be some of your

12 suggestions for that? So those are my three quick

13 questions?

14   LOURDES CINTRON: To everybody?

15   COUNCIL MEMBER BREWER: Yes whoever.

16   SARA LAPLANTE: [off mic]So I just want

17 to speak quickly to the formatting for the NYPD

18 data specifically. For example things like the

19 stop, question, frisk database. That is already in

20 a raw format that's easily digestible and oh sorry.

21   SARA LAPLANTE: It's in an open format

22 already. Well I shouldn't say that. It's in a data

23 format already that they publish on their website

24 and it, the issue there is really just that they're

25 not reporting it to the web portal. But those types

2 of things it takes maybe three clicks to put it

3 into a format that's open. And certainly any of

4 their comp stat data I would imagine would be

5 easily converted into a CSV file or something

6 that's, that's more easy to manipulate than, than a

7 PDF.

8   MATT BISHOP: You know I, I think you

9 know the, I'm not going to speak to any one

10 particular agency. I'm just going to speak

11 generally about being in… You want the data to be

12 all readable by apps because I think one of the

13 things to remember is not just you know what any

14 one analysis or journalist will bring together. But

15 you know some, sometimes the, these civic hackers

16 will, they'll come up with ways of playing with

17 this data that you know we just, we just never, we

18 never really thought of you know? Ways of doing

19 analysis or bringing things together in a way that,

20 so it needs to be readable by apps so JSON format

21 CSV.

22   CHAIRPERSON CABRERA: Alright thank you

23 so much and let me just take the moment as we close

24 here. I want to thank all of the groups that came

25 this to be honest with you, this was very

COMMITTEE ON TECHNOLOGY   116

2 informative, helpful, and we're going to move,

3 we're going to be purposed driven here, moving

4 forward here. And also I want to take a moment to

5 thank council member Brewer for really being the

6 Vanguard and the flag carrier for the open data

7 law. I don't, I was sitting here, I don't even

8 think none of us here what's going to be the

9 multiplier effect of this information that is so,

10 so neat and I saw in the hearing here the different

11 groups it just you know. It's opened up my mind as

12 to possibilities. But possibilities always bring

13 forth more possibilities. So I want to thank you. I

14 wanted to do that publically.

15   [applause]

16   CHAIRPERSON CABRERA: Commend this…

17   COUNCIL MEMBER BREWER: I want to thank

18 because it's a joint effort, this chair has been

19 very open. I'm very crazed on technology but I

20 think what's exciting about the testimony today

21 each testifier is phenomenal. I mean really really

22 high quality testimonies specific to the issue and

23 that's what's great about this community you're a

24 serious posse. Thank you.

25   [laughter]

1

2   [applause]

3   CHAIRPERSON CABRERA: Have a wonderful

4   day and again we are being joined by council member

5   Tish James. Enjoy your day. Thank you so much.

6   [gavel]

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

C E R T I F I C A T E

World Wide Dictation certifies that the

foregoing transcript is a true and accurate

record of the proceedings. We further certify that

there is no relation to any of the parties to

this action by blood or marriage, and that there

is interest in the outcome of this matter.

Date _____ November 04, 2013_____