Testimony of Gale A. Brewer, Manhattan Borough President
Regarding the Commission on Public Information and Communication
New York City Council Committee on Technology
February 12, 2019

My name is Gale Brewer, and I am the Manhattan Borough President. I thank Chairman Koo and the members of the Committee on Technology for scheduling this hearing. As you may know, I was a member of the Commission of Public Information and Communication (COPIC) representing the City Council during my tenure in the Council, and the experience was partly responsible for my being the prime sponsor of New York City's Open Data Law.

The desire to improve government transparency that led to the creation of COPIC in 1989 remains an important motivating force today. To that end, New York City has made great strides and now boasts the most robust open data offerings of any municipality in the United States. This was made possible through strong leadership and the important input of our civic hacker community.

But the Commission is in dire need of restructuring to remain relevant and fulfill its purpose. Section 1061 of the New York City Charter clearly outlines the various duties COPIC is supposed to undertake. Many of those duties have been left by the wayside, only to be picked up by other stakeholders. COPIC's annual public hearing on city information policies has been replaced by the New York City Council Committee on Technology's annual oversight hearing on the Open Data Portal. The annual report the Commission is supposed to publish is instead put together by the Department of Information Technology and Telecommunications (DOITT) and the Mayor's Office of Data Analytics (MODA). The charter tasks COPIC with making "recommendations regarding the application of new communications technology to improve public access to city produced or maintained information" [Section 1061 d (6)]. Instead, we lean on civic hack-tivists for this crucial input.

Beyond these role reassignments, COPIC is not living up to its original purpose in key areas. To the best of my knowledge, the Commission has not met since 2016 when the members helped to institute webcasts for public meetings. That was an achievement to be proud of. However, other ideas for civic tech could have been developed if the members of the Commission had met on a regular basis.

I said earlier that COPIC hasn't met since 2016 "to the best of my knowledge". That's because the Commission no longer has a website or a clear point of contact for information; this is ironic, considering its mandate for transparency.

Some of the reductions in the scope of COPIC resulted from a duplication of efforts among several agencies. However, there is clear value in having an organization composed of stakeholders from varied backgrounds dedicated to preserving government transparency. COPIC should be restructured and resourced to develop strategies to safeguard our open data platforms and the philosophy of open access. The restrictions on public information and its distortion and misuse for political purposes at the Federal level should alert us to the dangers that could occur locally.

The current degree of open data and access to government information was inconceivable in 1989. Looking to the future, we must reimagine the structure and role of COPIC in our open data ecosystem. It should be funded for a functioning website to ensure the public remains informed on data issues. The Commission should have a full-time staff that can help inform and enact the vision of the members while working in concert with DOITT and MODA to ensure we do not duplicate the efforts of city agencies. COPIC should meet quarterly to ensure a steady flow of information between internal and external stakeholders.

When Mayor de Blasio was the Public Advocate, COPIC met very rarely. I was the instigator of getting it to meet at all. As Public Advocate, de Blasio felt that any meaningful activity and agenda items that COPIC initiated would not be able to be implemented if staff funding was not in place. The Public Advocate's office did not have adequate funding to staff COPIC. I do not know why COPIC did not meet regularly more recently; the lack of funding for staff may have also been the reason. I was on the staff of Public Advocate Mark Green when COPIC was first initiated and there were regular meetings but even then a staff of 8 was contemplated and we did not have enough budget support.

Government transparency is vital. It makes government more accountable, empowers citizens and small businesses, and improves city services. There is much progress to be made on this front, and I will do all I can to ensure that the vision of 2012 continues to inspire and inform such initiatives, and that New York City remains a national leader in municipal data innovation. COPIC is part of that vision and must be reinvented to ensure the continued success of New York City's Open Data offerings.

Testimony by Andrew Rasiej, CEO and Founder Civic Hall
February 12, 2019
NYC Council Hearing on COPIC

Thank you for inviting me today to discuss a topic of vital importance to the future of our great city.

In 2005 I ran for the office of the Public Advocate on a platform to transform the office into a network of public advocates using Wifi and broadband technologies to connect with each other, build coalitions, and advocate for their communities. Very few people understood what I was talking about or believed it could be done. The NYTimes Editorial Board in my endorsement interview asked me to explain the term Wifi to them. Journalists covering the campaign asked me how I was possibly going to wire up the entire city on the Public Advocates measly budget. Candidates for Mayor the same year, asked me if Wi-Fi meant we would have to dig up the streets? Their doubt and ignorance was not so hard to understand. At the time. Social media was still a nascent emerging technology. There was no Twitter, no LinkedIn, and Facebook was still mostly a platform for college students. We carrying flip phones like a Mottorla StarTac and paying 50 cents for each text message we would send or receive and digital cameras were just beginning to become affordable for the masses. How things have changed.

Today we live in a hyper-connected world. The Internet and mobile technologies have transformed our economy and lives. Mayor de Blasio has ~~identified~~ made universal access to free Wi-Fi and low-cost broadband for all New Yorkers, a major policy goal. Our StarTac's have evolved into smartphones that for better or for worse, have become indispensable in how we live, work, and play. Every business not in technology is now rushing to transform into digital enterprises in order to compete with millions of tech start-ups looking to disrupt their marketplaces. Cloud computing has become ubiquitous and almost free. Students are choosing careers in data science in droves. Professionals in every major industry are taking classes in digital skills, to help them either perform better at their jobs or give them the tools they need to start their own startups.

And although there are serious issues and challenges that all of these technologies present, particularly around the use of and safety of people's private data, the distribution of fake and biased media, as well as an ever-growing threat of cyber warfare, the technology itself continues to evolve dramatically and an ever faster rate. Artificial Intelligence is now being embedded in every new technology. 5G networks and small cell technologies are being deployed, and data science is transforming commerce, medicine, and finance.

However, one place technology has failed to impact is the government.
Walk into many New York City government offices and you will still see papers piled on top of file cabinets. Go online to apply for government service and you will most likely see the same website that was built 15 years ago. Our public schools and our police precincts essentially look and operate the same way they did 40 years ago. We can order

a special meal to be delivered with a few swipes of a smartphone but applying for Snap benefits still requires a paper application faxed to a government office.

That's not to say that the government hasn't made any progress in the use of technology. The NYC Open Data Law championed by then City Councilwoman Gale Brewer, has opened up vast amounts of data for public consumption offering more transparency and catalyzing the creating of many new applications built by the private sector. Our subways sometimes tell us when trains are arriving, and EZ pass has essentially eliminated lines to pay tolls at our bridges and tunnels. Our schools are now wired with broadband, and thanks to funding by the Manhattan DA, Police officers now carry smartphones too. But unfortunately, such innovations are few and far between. City agencies do not have procedures to standardized data, there are few resources to train city workers in digital skills to help them do their jobs, much less to learn critical cybersecurity procedures to keep our city safe. The city still maintains much of its data in siloed physical data centers wasting millions of dollars on physical infrastructure that is also costly to maintain and to power. While the rest of the world is designing and driving Teslas, NYC government is still driving a 1985 Pontiac.

It doesn't have to be this way. New technologies offer New York City opportunities to transform our city into a more efficient, effective, and responsive. I am sure everyone in this room and everyone who isn't here and cares about our city and its people would agree with me. The question is how to make it happen?

One way to start is for New York City government to develop a strategic plan to transform itself into a more modern and digitally capable institution. In the past, the city has relied on the Department of Information Technology and Telecommunications (DoITT) for about half of its agencies technology needs. The problem is that the other agencies DoITT does not service, each has its own technology infrastructure and procedures. Few are coordinated with each other, share data, or outside services. Virtually none provide professional development for their staff. There is no single agency that is responsible for the whole city's technology infrastructure or capacity.

Also, because the people working at these agencies are busy running them every day, their ability to develop a strategic plan for technology much less implement one is very limited. Therefore it is now imperative that COPEC and the City Council create a new and separate strategic planning commission which would include representatives from the offices of the Mayor, Comptroller, and Public Advocate, along with advisors from the private sector to research what the city can do to upgrade its proficiency in the use of information technology. If the goal of this hearing is to discuss ways the city could increase government transparency, the public's access to government information, protecting the privacy of New Yorkers, and facilitate the sharing of data by government agencies, the pre-requisite is for the city to develop just such a plan. As well-meaning and thoughtful the other recommendations being offered today may be, they will be more likely to succeed if a strategic plan was developed particularly with a non-political perspective, and the support of private sector and other advisers.

Today, if I was the Public Advocate, (by the way I am not running again), the idea of building a network of all of the cities public advocates connected by Wi-Fi and broadband would not only be believable but it would also be possible.

What our city needs most of all when facing the challenges and the opportunities of the inevitable technological future are having leaders who have a vision and a will to change the way the city works.

If not now when? If not you, who?

Andrew Rasiej
Founder and CEO, Civic Hall.

To: Acting Public Advocate and Chair of the Technology Committee
From: Noel Hidalgo, Executive Director of BetaNYC
Re: Oversight - Commission on Public Information and Communication's
Collaboration in Developing City Information Policies and Promoting Governmental
Transparency

β

Tuesday, 12 February 2019

Dear Acting Public Advocate and New York City Council Speaker Johnson,
Committee Chair Koo, and Technology Committee Members

It is an honor to have this opportunity to represent New York City's civic technology, design, and data community.

I am the Executive Director of BetaNYC, a member driven good government non-profit organization. *We are advocates for a City government that is for the people, by the people, and for the digital era.*

In 2009, a group of neighbors started meeting to discuss the future of municipal open data and technology because they were concerned about a lack of open data and expensive technology procurements.

Over the last ten years, our 5,100+ members have fought to improve people's lives through technology, data and design.[1] We have watched the past three Public Advocates apoint COPIC members, host one meeting per term, and walk out of the office with little accomplishment. We have watched every Public Advocate publish flowery press releases only for them to disappear like tears in rain.

In 2012, we joined with Council Member Gale A Brewer to support the City's open data law. COPIC's absence is why we fought for the City's open data law.

In 2014, we published a *People's Roadmap to a Digital New York City*[2]. It outlined how our City could adopt modern, agile practices to meet pressing needs for a more efficient, participatory, and transparent government. Additionally, we proposed 34 ideas that resulted in the following transformative legislation:
- placing the City Record online and in a machine readable format,
- ensuring that our Charter and laws are owned by the people, not a corporation,
- strengthening the NYC's open data laws through seven interlocking pieces of legislation,
- Formalizing the City's Chief Analytics Officer and the Mayor's Office of Data Analytics into the charter.

---

[1] https://www.meetup.com/betanyc/
[2] http://nycroadmap.us

The *People's Roadmap* outlined ideas that required government partnership. And, for the past four years, we have worked successfully with the Manhattan Borough President Gale Brewer, Brooklyn Borough President Eric Adams, Mayor's Office of Data Analytics, Manhattan and Brooklyn Community Boards, CUNY Service Corps, the Fund for the City of New York, and the Alfred P. Sloan Foundation to study and test how communication technologies and open data can equip the public to improve their decision making. We have published three reports and filed numerous data set enhancements with the goal of improving community decision making.

Here is a list of a few achievements we have made; we have created the nation's first FREE municipal open data bootcamp[3] — suggestions on how community board could better use communication technologies, including their websites (which DOITT is working on)[4] — we convinced DOITT to be a part of the district needs process — we've educated, mentored, and employed over 50 City University of New York undergraduates[5] (one of our Alumni is now one of your aides) — we've built a suite of specialized open data tools[6] — we've documented, in detail, how information flows through community board meetings and the data they need to improve decision making[7] — we've taught over one thousand New Yorkers how to put NYC's open data to use — we've enriched a local community of open data professionals and advocates by hosting three annual citywide open data festivals, with the fourth co-hosted with MODA[8] coming up on Saturday, 2 March.

We are New York City's biggest open data fan — we have partnered with NYC Parks, NYC 311, NYC Planning, the Mayor's Office of Data Analytics, and NYC BigApps to host numerous data jams and hackathons that explore and demystify the City's data.

We are working with Department of Education's Computer Science for All program, inspiring the next generation of tech-empowered citizens. Through five borough wide hackathons and a citywide hackathon, middle and high school students are seeing themselves within the fabric of this city and its data. These young leaders are using their computer science skills to express their voice and contribute to a better city.

This was done this because someone needed to be the shepherd.

With the unflinching support of Manhattan Borough President Gale Brewer and Mayor's Office of Data Analytics, we have marshaled private funds and community engagement to outline improvements to the City's data and community technology infrastructure

---

[3] https://beta.nyc/products/nyc-opendata-classes/
[4] https://beta.nyc/publications/betanycs-civic-innovation-fellows-community-board-technology-needs-report-2018/
[5] https://beta.nyc/programs/civic-innovation-lab/
[6] https://beta.nyc/products/
[7] https://beta.nyc/publications/
[8] https://schoolofdata.nyc

In testimony to the 2018 Charter Revision Commission, BetaNYC believes that COPIC needs to be reconceived for the 21st Century.[9]

At its core, COPIC functions to address three fundamental issues:
- **Oversight of government using existing communication technologies**
- **Access to public government information and data**
- **Government adoption of new communication technologies**

At this time, we firmly believe that COPIC's mandate is being executed though the collection of open data laws and reporting requirements.[10] Council Members and the Technology Committee Chair have been extremely supportive and responsive to oversight hearings and community requests. This being said, there is an explicit need for COPIC's oversight functions.

If there is ever a time for communication technology leadership, it is now.

We've had an acting Chief Technology Officer[11] for as long as we were missing a Chief Analytics Officer [12]. We've spent the last year working with DOITT to modernize community boards' websites with their staff. For the past three years, we have observed and documented an explicit need for a contact management system for community boards. For the past four years, the city has told me that it will insource and adopt agile processes and technologies in line with what we have outlined. Yet, all that has come out of this are agency teams fighting night and day to implement those practices.[13][14]

The underlying issues that COPIC is attempting to address are just as real today as they were 30 years ago!

The issues you have with 311 are fundamentally about poorly designed workflows— workflows that are enshrined in, rather than making use of, technology. The issues you have about after-school programs exist because people don't know how to use city databases — they are not "user friendly and — they don't feature proper data validation. You raise issues about community board websites being outdated — without recognizing the inherent problem that they are supported by a tool built in the last century. If we want COPIC to exist and thrive, we have to take a hard look at what they oversee.

As of this month, most people access federal government websites with their mobile or tablet devices.[15] Does New York City know this? No, we don't keep track of these metrics, and thus we don't build for mobile, which means, people have trouble accessing government sites. What happened to our Chief Digital Officer? What is going on with alpha.nyc.gov? What is DOITT doing to address government services for the digital era? Where are the marshaling of municipal resources to improve our City's communications technology?

---

[9] https://beta.nyc/?s=COPIC

[10] https://opendata.cityofnewyork.us/open-data-law/

[11] http://www.govtech.com/people/New-York-City-CTO-Miguel-Gamino-Departs-for-Private-Sector.html

[12] http://www.gothamgazette.com/city/8001-de-blasio-names-new-chief-analytics-officer

[13] https://planninglabs.nyc/

[14] https://www1.nyc.gov/site/opportunity/portfolio/service-design-studio.page

[15] https://twitter.com/gbinal/status/1092862863587373056

We are the world's greatest city and one of the largest. We need to consider the implications of weak digital leadership. Even with the strongest COPIC, its impact would be a mear pinprick on what is needed. New York City needs a properly resourced government body that ensures city agencies are building / adopting modern, open technology. These teams should reflect other digital government success like 18F[16] and the United States Digital Service[17].

Beside my ardent belief in building technology from within government, I question how the Public Advocate, COPIC, and the Mayor's Offices will exist in a world with limited resources and expertise.

I conclude my testimony with a list of questions to help guide you through the reformation of COPIC.

Fundamental Questions:
- What is the future of public oversight of government data and public information?
- How do we ensure that COPIC's powers are complementary and do not creative duplicative work for powers now enshrined in the City's collection of open data?
- What is the role of the Public Advocate in guiding and setting technology and data standards?
- Where are the archives of COPIC's advisory decisions? Have they been implemented? Have they been useful? What is COPIC's impact?

Questions for today:
- What has come of COPIC's current board? What is COPIC's current staffing, agenda, and resources?
- What is the role of DOITT and its Commissioner, the Chief Information Officer, in providing public facing digital services within NYC?
- What is the role of Chief Technology Officer or the Chief Digital Officer in providing public facing digital services within NYC?
- What is the role of the Public Advocate overseeing technology contracts and the City's information technology services?

Long term questions for COPIC:
- How do we get real public representation on COPIC? (Out of 12 members, four are paid and two are are not appointed by the Mayor or City Council.)
- What is the role of the Public Advocate within the City's production of technology and data?
- How can we put adequate resources and oversight in place to bring NYC's public information platforms into the 21st century?
- In the face of an Administration that would want to dismantle / un-fund existing open data programs, what systems can we put in place to ensure appropriate oversight and protect the future of open data and public information technology platforms?

[16] https://18f.gsa.gov/
[17] https://www.usds.gov/

STATEMENT OF
ALBERT FOX CAHN, ESQ.
EXECUTIVE DIRECTOR
SURVEILLANCE TECHNOLOGY OVERSIGHT PROJECT, INC.

BEFORE THE
COMMITTEE ON TECHNOLOGY
NEW YORK CITY COUNCIL

FOR A HEARING CONCERNING,
OVERSIGHT - COMMISSION ON PUBLIC INFORMATION AND
COMMUNICATION'S COLLABORATION IN DEVELOPING CITY INFORMATION
POLICIES AND PROMOTING GOVERNMENTAL TRANSPARENCY.

PRESENTED
February 12, 2019

Good morning, my name is Albert Fox Cahn, and I serve as the Executive Director for the Surveillance Technology Oversight Project ("STOP"). STOP advocates and litigates the privacy rights of New Yorkers impacted by suspicionless, warrantless surveillance. I commend the committee and Chairman Koo for today's hearing; for bringing much-needed attention to the fight for governmental transparency.

For the past year, I've have been proud to partner with the City as part of its Automated Decision Systems Task Force, meeting with City leaders, academics, and advocates to shape recommendations for the future role of artificial intelligence in New York City Government. As part of my role in the task force, I have noted that while Transparency is crucial in every area of government, it is nowhere more vital than in policing, where mistakes can quickly rob New Yorkers of their liberty, or even their life.

As part of today's hearing, I urge the committee to note the urgent need for greater transparency of New York City Police Department ("NYPD") surveillance practices, especially those tools that use artificial intelligence and other automated decision systems. Specifically, I speak today in support of the POST Act[1], which would be an important step forward in strengthening police oversight, promoting public safety, and safeguarding New Yorkers' privacy rights.

Historically, the NYPD deployed novel and highly invasive surveillance technologies in ways that circumvented democratic oversight and accountability. The NYPD used private and federal funds, without any disclosure to the lawmakers we depend-on to oversee our police forces. With this unaccountable funding, the NYPD was able to deploy tools like "stingrays," fake cell towers that collect sensitive location and communications data.[2] Like many of the NYPD's new tools, stingrays spy not only on the target of an investigation, but also on untold numbers of innocent bystanders.[3]

Let me be clear, the POST Act does not prohibit the NYPD from using new surveillance tools. Rather, it merely secures this Council's indispensable role in reviewing when and how such tools are deployed. Under the POST Act, the NYPD must issue an "impact and use policy" report when choosing to use a new surveillance tool.[4] This report must describe the technology, rules, and guidelines for the use of that technology, and safeguards for protecting any data collected.[5] The City

---

[1] Public Oversight of Surveillance Technology (POST) Act, Int 0487-2018.

[2] Joseph Goldstein, *New York Police Are Using Covert Cellphone Trackers, Civil Liberties Group Says*, N.Y. TIMES, Feb. 11, 2016, https://www.nytimes.com/2016/02/12/nyregion/new-york-police-dept-cellphone-tracking-stingrays.html.

[3] *Id.*

[4] N.Y. CITY COUNCIL 1482 § 1 (N.Y. 2017), ch. 1, 14 ADMIN. CODE OF N.Y.C. § 14-167(b) (as proposed)

[5] *Id.* at 14-167(a) (as proposed)

Council and the people of New York City would then be allowed to provide feedback on such an acquisition.[6] Thus, the POST Act strikes a delicate balance, requiring sufficient information to ensure oversight, while protecting operational details, sources, and methods.

Civilian oversight of policing and intelligence gathering is not only a fundamental American value, it is essential for effective policing. As then-President Obama's Task Force on 21st Century Policing found, "[l]aw enforcement agencies should establish a culture of transparency and accountability in order to build public trust and legitimacy."[7] The NYPD's current procurement methods are not only undemocratic, but they harm the NYPD's very mission of promoting public safety.

The POST Act will benefit all New Yorkers, but it will offer particularly powerful protection for our Muslim neighbours. For years, Muslim New Yorkers have faced a pattern of unjust and unconstitutional NYPD surveillance. Specifically, the NYPD's Intelligence Division engaged in extensive, suspicionless surveillance of majority Muslim neighbourhoods and Muslim families.[8] Additionally, NYPD officials have conducted blanket surveillance of entire mosques, surveilling men, women, and children for nothing more than practicing their faith.[9] Some local businesses have even been classified as "place[s] of concern" for nothing more than having customers of middle eastern dissent.[10]

The facts are clear, the NYPD is engaged in widespread, discriminatory policing practices. According to the Office of the Inspector General for the NYPD ("OIG"), over 95% of recent NYPD political and religious investigations targeted Muslim individuals and organizations.[11] The

---

[6] *Id.* at 14-167(e-f) (as proposed)

[7] PRESIDENT'S TASK FORCE ON 21ST CENTURY POLICING, FINAL REPORT OF THE PRESIDENT'S TASK FORCE ON 21ST CENTURY POLICING 12 (2015), https://cops.usdoj.gov/pdf/taskforce/taskforce_finalreport.pdf.

[8] Matt Apuzzo & Joseph Goldstein, *New York Drops Unit That Spied on Muslims*, N.Y. TIMES, Apr. 15, 2014, https://www.nytimes.com/2014/04/16/nyregion/police-unit-that-spied-on-muslims-is-disbanded.html?_r=0; *see also* DIALA SHAMAS & NERMEEN ARASTU, MUSLIM AM. CIVIL LIBERTIES COAL., CREATING LAW ENF'T ACCOUNTABILITY & RESPONSIBILITY & ASIAN AM. LEGAL DEF. & EDUC. FUND, MAPPING MUSLIMS: NYPD SPYING AND ITS IMPACT ON AMERICAN MUSLIMS 10 (2013), https://www.law.cuny.edu/academics/clinics/immigration/clear/Mapping-Muslims.pdf.

[9] Apuzzo & Goldstein, *supra* note 7.

[10] Adam Goldman & Matt Apuzzo, *NYPD: Muslim Spying Led to No Leads, Terror Cases*, ASSOCIATED PRESS, Aug. 21, 2012, https://www.ap.org/ap-in-the-news/2012/nypd-muslim-spying-led-to-no-leads-terror-cases.

[11] OFFICE OF THE INSPECTOR GEN. FOR THE N.Y. POLICE DEP'T, N.Y. CITY DEP'T OF INVESTIGATION, AN INVESTIGATION OF NYPD'S COMPLIANCE WITH RULES GOVERNING INVESTIGATIONS OF POLITICAL ACTIVITY 1 n.1 (2016), https://www1.nyc.gov/assets/oignypd/downloads/pdf/oig_intel_report_823_final_for_release.pdf. In

pattern of discriminatory surveillance is completely at odds with the fact that the overwhelming majority of terrorist attacks in the United States are committed by right-wing extremists and white supremacists. Let me repeat that fact, since it is so often lost in our media environment: right-wing extremists and white supremacists commit the overwhelming majority of terrorist attacks in the United States. That is not the findings of partisans, that is the conclusion of groups ranging from the Anti-Defamation League, to the Southern Poverty Law Center, to the U.S. General Accountability Office.[12]

In contrast to the undercover practices documented above, the novel NYPD surveillance practices governed by the POST Act often are completely invisible to the target, making them much more dangerous to our freedom of speech and religion. The need for oversight is only heightened by the NYPD's clear track record of disregarding those few existing restrictions on surveillance of protected First Amendment activity. According to the OIG, over half of NYPD intelligence investigations continued even after the legal authorization for them expired.[13] Also, the OIG found that the NYPD frequently violated legal guidelines governing these investigations in other ways, such as through its use of boilerplate language in undercover officer authorization forms.[14]

In light of the foregoing, we urge this City Council to enact the POST Act. This legislation will provide vital transparency for the NYPD's acquisition of, and use of, surveillance technology. I thank you for giving me the opportunity to address these urgent issues, and I look forward to working with the Council to safeguard the rights of Muslim New Yorkers in the months and years to come.

(Cont'd from preceding page)

its investigation, the OIG reviewed a random selection of 20% of cases closed or discontinued between 2010 and 2015 of each case type. *Id.* at 14.

[12] U.S. Gov't Accountability Office, GAO-17-300, Countering Violent Extremism: Actions Needed to Define Strategy and Assess Progress of Federal Efforts 4 (2017), https://www.gao.gov/assets/690/683984.pdf; David Neiwert, *Trump's Second Travel Ban Once Again Misidentifies Source of Domestic Terrorist Threat*, Southern Poverty Law Center (Mar. 13, 2017), https://www.splcenter.org/hatewatch/2017/03/13/trumps-second-travel-ban-once-again-misidentifies-source-domestic-terrorist-threat; *Murder and Extremism in the United States in 2016*, Anti-Defamation League, https://www.adl.org/education/resources/reports/murder-and-extremism-in-the-united-states-in-2016 (last visited June 13, 2017).

[13] Office of the Inspector Gen. for the N.Y. Police Dep't, *supra* note 25, at 1.

[14] *Id.* Such conduct undermines the ability of independent bodies to effectively review police compliance with legal guidelines. *Id.* at 2.

CENTER ON CIVIL JUSTICE
at NYU School of Law

Testimony by David Siffert, Research Coordinator, Center on Civil Justice at NYU School of
Law
Joint Hearing of the Commission on Public Information and Communication (COPIC) and the
Technology Committee of the New York City Council
250 Broadway, 14th floor Committee Room
February 12, 2019

Thank you for inviting me here to speak today. I am the Research Coordinator at the Center on
Civil Justice at NYU School of Law. The Center is dedicated to the study of the civil justice
system in the United States and how it can continue to fulfill its purposes. The Center is directed
by Peter Zimroth, former Corporation Counsel for New York City. Its faculty co-directors are
Arthur Miller, Samuel Issacharoff, Troy McKenzie, and Geoffrey Miller, some of the most
distinguished law professors in the country. Our Board of Advisers is chaired by Sheila
Birnbaum, one of the top trial lawyers in America, and contains some of the nation's most
accomplished plaintiffs lawyers, defense lawyers, and judges.

One major focus of the Center is providing access to data and information. One of our projects
is a simple, online, searchable document database on the subject of third party litigation funding.
The website, which we are calling the Dispute Financing Library, will serve as a neutral, quality
repository for the collection of information and data about this new, growing, and largely
unknown industry.

Another project we are working on involves working with judges to help them issue orders to
administrative agents in large cases to make public the data they collect in those cases. This
project involves balancing some of the privacy concerns that are being discussed at this hearing
and making sure that the information we collect is not only anonymized but also cannot be
"reverse-engineered" to identify any of the individual claimants in these cases.

Most recently, the Center hosted a conference on Artificial Intelligence in a Democratic Society.
The conference discussed the use of data, algorithms, and machine learning, and how to make
sure that AI technologies are developed responsibly.

As a result of this work, the Center is aware of a Scylla and Charybdis facing the worlds of law
and government – we are far behind-the-ball in terms of creating, distributing, and using data to
make our government and legal system work better, but we also lack the institutional protections
and technical know-how to ensure that this data is used responsibly, in a way that will protect
citizens' privacy and ensure that data biases are minimized.

Many of the speakers here can better explain the excellent uses for data in government and our legal system, or about the dangers to our privacy, or about the risks of algorithmic bias. Some may have answers and policy prescriptions for how to use data, how not to use data, and how to regulate the use of data. I would simply like to highlight one thing that I believe can help navigate the Scylla and Charybdis we are facing – and that is Education.

The Center is currently working on educational programs for lawyers and judges to teach them about data, algorithms, and automated decision-making. We hope to teach what is available, what is not available, what is appropriate to use in what contexts, and how to use it responsibly.

But the use of data in New York City extends far beyond the legal profession. The City needs to make substantial efforts to increase statistical literacy across-the-board. Of course, those in City Government who already handle data need to understand concerns about privacy and bias. Those in City Government who don't already work with data need to have sufficient understanding of what data exists to know whether that data could be used to improve the work they do.

But education cannot start and end within the government. Students in New York City's public schools need to have baseline levels of statistical competence if they are going to compete in the modern economy. We also need to create a talent pool that can lead us to a New York that has true, responsible adoption of data use. We need a populace that understands data – what it is, what it can do, how to use it appropriately, and the dangers of using it inappropriately. The Center on Civil Justice and New York University are working hard to supply this education to New Yorkers, but we need your help.

Thank you very much for your time.

Written Testimony of Angel Diaz
Counsel, Liberty & National Security Program
Brennan Center for Justice *at NYU School of Law*
Before the
New York City Council Committee on Technology
Regarding
Oversight – Commission on Public Information and Communication's
Collaboration in Developing City Information Policies
and Promoting Governmental Transparency
February 12, 2019

Good afternoon Chairman Koo and members of the Committee on Technology. My name is Angel Diaz, and I am Counsel to the Liberty and National Security Program at the Brennan Center for Justice. I am pleased to be testifying today about how the Commission on Public Information and Communication (COPIC) can help advance policies that increase governmental transparency.

The Brennan Center is a nonpartisan law and policy institute that seeks to improve our systems of democracy and justice. The Liberty and National Security Program focuses on restoring the proper flow of information between the government and the people by securing increased public access to public information; ensuring government policies targeting suspected criminals and terrorists do so effectively and without religious or ethnic profiling; and securing appropriate government oversight and accountability.

As part of this work, we actively seek greater transparency and oversight of the NYPD's surveillance tools. The NYPD has touted itself as being the most transparent police department in the world.[1] But in fact, the NYPD has frequently resisted transparency, requiring lawyers, journalists, and others to expend significant resources in order to obtain even basic information that is of critical interest to the public.[2]

For example, the Brennan Center is party to a multi-year legal dispute with the NYPD to obtain information about the Department's use of predictive policing technologies. These systems rely on

---

[1] *See* JPat Brown, Five Examples of the NYPD's Commitment to "Transparency," MUCKROCK (June 14, 2017), https://www.muckrock.com/news/archives/2017/jun/14/five-examples-nypd-transparency/.
[2] *See*, e.g., Adam Klasfield, Sound-Cannon Case Heralds E-Transparency for NYPD, COURTHOUSE NEWS (June 30, 2017), https://www.courthousenews.com/sound-cannons-case-heralds-e-transparency-nypd/; Brown, supra note 9.

algorithms to analyze large data sets and generate statistical estimates about crime, which are used to direct police resources.

Predictive policing tools have been roundly criticized by civil rights and civil liberties advocates,[3] as they often rely on historic crime data that both reflects and recreates decades of biased enforcement against communities of color.[4] In addition, there is little consensus that predictive policing is actually effective in predicting and reducing crime.[5] There is a common refrain that predictive policing predicts policing – it does not predict crime.[6]

Despite these concerns, former Police Commissioner Bratton and Mayor de Blasio announced in 2016 that the NYPD planned to spend $45 million on predictive policing technologies over the next 5 years.[7] We believed it was critical for the public to know more about the Department's existing systems, as well as any future versions of it. We therefore filed a public records request in July 2016 for a range of documents that would shed light on the NYPD's predictive policing efforts, including information about what type of information was fed into these algorithms and the results they generated.

The NYPD denied our initial request and subsequent appeal, forcing the Brennan Center to file suit in late 2016.[8] Despite months of good faith negotiations, the NYPD kept stonewalling, refusing to produce most of the documents we requested. In late 2017, a judge finally ordered the Department to produce records about its testing, development, and use of predictive policing tools.[9] But even then,

---

[3] *See*, e.g., Leadership Conference on Civil and Human Rights, et al., Predictive Policing Today: A Shared Statement of Civil Rights Concerns (Aug. 31, 2016), available at
http://civilrightsdocs.info/pdf/FINAL_JointStatementPredictivePolicing.pdf.

[4] *See*, e.g., Jack Smith IV, Crime-prediction Tool PredPol Amplifies Racially Biased Policing, Study Shows, MIC (Oct. 9, 2016), https://mic.com/articles/156286/crime-predictiontool-pred-pol-only-amplifies-racially-biasedpolicing-study-shows (last visited Oct. 15, 2017); see also Laura Nahmias, NYPD Testing Crime-Forecast Software, POLITICO (July 8, 2015, 5:52 AM EDT), http://www.politico.com/states/new-york/cityhall/story/2015/07/nypd-testing-crime-forecast-software-090820 (quoting maker of predictive policing software as noting the importance of assessing "how we apply statistics and data in a way that's going to be sensitive to civil rights and surveillance and privacy concerns").

[5] *See*, e.g., WILLIAM J. HAYES, NAVAL POSTGRADUATE SCH., CASE STUDIES OF PREDICTIVE ANALYSIS APPLICATIONS IN LAW ENFORCEMENT (Dec. 2015), available at https://www.hsdl.org/?view&did=790324; Martin Maximino, The Effectiveness of Predictive Policing: Lessons From A Randomized Controlled Trial, JOURNALIST
RES. (last updated Nov. 6, 2014), https://journalistsresource.org/studies/government/criminaljustice/predictive-policing-randomized-controlled-trial; Matt Stroud, Chicago's Predictive Policing Tool Just Failed
A Major Test (Aug. 19, 2016, 10:28 AM EDT), https://www.theverge.com/2016/8/19/12552384/chicagoheat-list-tool-failed-rand-test.

[6] *See* Ezekiel Edwards, Predictive Policing Software Is More Accurate At Predicting Policing Than Predicting Crime, HUFFPOST (Aug. 31, 2016, 2:58 EDT), https://www.huffingtonpost.com/entry/predictive-policingreform_us_57c6ffe0e4b0e60d31dc9120.

[7] *See*, e.g., CITY OF N.Y., DEVELOPING THE NYPD'S INFORMATION TECHNOLOGY 6-7, available at
http://home.nyc.gov/html/nypd/html/home/POA/pdf/Technology.pdf (last visited Oct. 16, 2017); Mayor de Blasio, Police Commissioner Bratton Announce CompStat 2.0, CITY OF N.Y. (Feb. 23, 2016), http://www1.nyc.gov/office-of-the-mayor/news/199-16/transcript-mayor-deblasio-policecommissioner-bratton-compstat-2-0#/0 (last visited February 12, 2019).

[8] *See* Rachel Levinson-Waldman and Erica Posey, "Predictive Policing Goes to Court," September 5, 2017,
http://www.brennancenter.org/blog/predictive-policing-goes-court.

[9] *See* Rachel Levinson-Waldman and Erica Posey, "Court: Public Deserves to Know How NYPD Uses Predictive Policing Software," January 28, 2018, https://www.brennancenter.org/blog/court-rejects-nypd-attempts-shield-predictive-policing-disclosure.

it took almost a full year from the judge's order before the NYPD finally produced some of the information in our request.

This is just one example of the NYPD's many surveillance tools. The NYPD also has object-recognition software[10] that can identify individuals based on their skin tone; it deploys cell-site simulators[11] that can trick every phone in their vicinity into sharing identifying information; and it operates a Domain Awareness System[12] that combines information from NYPD records and databases with the thousands of public and private security cameras that blanket New York City. Earlier this year, a public records request showed how the NYPD was engaging in social media monitoring of Black Lives Matter activists during protests back in 2014.[13]

In short, there is a serious need for mandatory transparency and oversight when it comes to the NYPD, to ensure that the Department is disclosing the records and other data that the public is entitled to access. This is why the Brennan Center calls on this Committee and the COPIC to support the POST Act, a bill that was re-introduced by Council Member Gibson and Co-Sponsored by Council Member Lander of this Committee. The POST Act would require the NYPD to publicly report on the surveillance tools it uses and describe the rules it has for using them.[14]

Specifically, the bill would require the NYPD to create an "impact and use policy" for each surveillance tool it uses now or in the future. These reports would provide descriptions and capabilities of each technology, establish rules and guidelines for their use, and contain policies for retaining and using any data collected by a surveillance tool. The impact and use policy would also describe safeguards to protect the privacy of New Yorkers and outline any internal audit and oversight mechanisms.

Although the NYPD may not wish to discuss the surveillance tools they use, a strong local democracy like New York City requires at least a basic level of information about what its local police are doing and how they're doing it. The POST Act is carefully balanced to achieve transparency and accountability while avoiding the disclosure of operational details that might compromise police investigations or harm public safety.

In an increasingly data-driven society, it is important that our elected officials do not let transparency and accountability fall by the wayside. We commend this Committee for addressing this important issue and urge you to support measures that empower the public to hold their elected officials accountable.

Thank you again for the opportunity to testify today. I am happy to answer any questions.
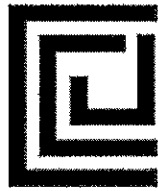
---

[10] *See*, George Joseph and Kenneth Lipp, "IBM Used NYPD Surveillance Footage to Develop Technology That Lets Police Search By Skin Color," The Intercept, September 6, 2018, https://theintercept.com/2018/09/06/nypd-surveillance-camera-skin-tone-search/.

[11] *See* https://www.aclu.org/issues/privacy-technology/surveillance-technologies/stingray-tracking-devices.

[12] Joe Coscarelli, "The NYPD's Domain Awareness System Is Watching You," New York Magazine, August 12, 2012, http://nymag.com/daily/intelligencer/2012/08/nypd-domain-awareness-system-microsoftis-watching-you.html.

[13] *See* George Joseph, "Years After Protests, NYPD Retains Photos of Black Lives Matter Activists," The Appeal, January 17, 2019, https://theappeal.org/years-after-protests-nypd-retains-photos-of-black-lives-matter-activists/.

[14] For more on the POST Act, short for Public Oversight of Surveillance Technology Act, see "The Public Oversight of Surveillance Technology (POST) Act: A Resource Page, available at https://www.brennancenter.org/analysis/public-oversight-surveillance-technology-post-act-resource-page.

CENTER ON CIVIL JUSTICE
at NYU School of Law

Testimony by David Siffert, Research Coordinator, Center on Civil Justice at NYU School of
Law
Joint Hearing of the Commission on Public Information and Communication (COPIC) and the
Technology Committee of the New York City Council
250 Broadway, 14th floor Committee Room
February 12, 2019

Thank you for inviting me here to speak today. I am the Research Coordinator at the Center on
Civil Justice at NYU School of Law. The Center is dedicated to the study of the civil justice
system in the United States and how it can continue to fulfill its purposes. The Center is directed
by Peter Zimroth, former Corporation Counsel for New York City. Its faculty co-directors are
Arthur Miller, Samuel Issacharoff, Troy McKenzie, and Geoffrey Miller, some of the most
distinguished law professors in the country. Our Board of Advisers is chaired by Sheila
Birnbaum, one of the top trial lawyers in America, and contains some of the nation's most
accomplished plaintiffs lawyers, defense lawyers, and judges.

One major focus of the Center is providing access to data and information. One of our projects
is a simple, online, searchable document database on the subject of third party litigation funding.
The website, which we are calling the Dispute Financing Library, will serve as a neutral, quality
repository for the collection of information and data about this new, growing, and largely
unknown industry.

Another project we are working on involves working with judges to help them issue orders to
administrative agents in large cases to make public the data they collect in those cases. This
project involves balancing some of the privacy concerns that are being discussed at this hearing
and making sure that the information we collect is not only anonymized but also cannot be
"reverse-engineered" to identify any of the individual claimants in these cases.

Most recently, the Center hosted a conference on Artificial Intelligence in a Democratic Society.
The conference discussed the use of data, algorithms, and machine learning, and how to make
sure that AI technologies are developed responsibly.

As a result of this work, the Center is aware of a Scylla and Charybdis facing the worlds of law
and government – we are far behind-the-ball in terms of creating, distributing, and using data to
make our government and legal system work better, but we also lack the institutional protections
and technical know-how to ensure that this data is used responsibly, in a way that will protect
citizens' privacy and ensure that data biases are minimized.

Many of the speakers here can better explain the excellent uses for data in government and our legal system, or about the dangers to our privacy, or about the risks of algorithmic bias. Some may have answers and policy prescriptions for how to use data, how not to use data, and how to regulate the use of data. I would simply like to highlight one thing that I believe can help navigate the Scylla and Charybdis we are facing – and that is Education.

The Center is currently working on educational programs for lawyers and judges to teach them about data, algorithms, and automated decision-making. We hope to teach what is available, what is not available, what is appropriate to use in what contexts, and how to use it responsibly.

But the use of data in New York City extends far beyond the legal profession. The City needs to make substantial efforts to increase statistical literacy across-the-board. Of course, those in City Government who already handle data need to understand concerns about privacy and bias. Those in City Government who don't already work with data need to have sufficient understanding of what data exists to know whether that data could be used to improve the work they do.

But education cannot start and end within the government. Students in New York City's public schools need to have baseline levels of statistical competence if they are going to compete in the modern economy. We also need to create a talent pool that can lead us to a New York that has true, responsible adoption of data use. We need a populace that understands data – what it is, what it can do, how to use it appropriately, and the dangers of using it inappropriately. The Center on Civil Justice and New York University are working hard to supply this education to New Yorkers, but we need your help.

Thank you very much for your time.

NYU | WAGNER

JULIA LANE
Professor, Wagner School
Professor at NYU Center for Urban
Science and Progress
NYU Provostial Fellow for Innovation
Analytics

Robert F. Wagner Graduate School of
Public Service
New York University
295 Lafayette Street, 2nd Floor
New York, NY 10012

P: 646 208 4990

julia.lane@nyu.edu

Dear Chair Koo, Acting Public Advocate Johnson, and members of the committee and commission:

Our world has been transformed over the past twenty years. We're living in a new Industrial Revolution. Data are automating the way in which we do everything – live, work, shop and play. The private sector uses data to figure out what consumers want now, and uses data to deliver to them – with almost frightening efficiency.

Except government agencies that produce data for decision-making. Those agencies are often doing their job in the data dark ages – producing information manually, and using dated tools and methods to do so. But governments can't simply turn to the private sector to do their job for them. Government data are used to allocate millions of taxpayer dollars, affect millions of citizens, and are intended to serve the public, not generate profit. Governments need to be trustworthy, generate information that are measured well and consistently over time and need to ensure that confidential information is protected.

The bad news is that the way governments produce data won't change by itself. In the private sector, market forces create the impetus for change, because organizations that don't adapt get driven out of business. There's no similar force driving government change. Over the past 30 years, I've worked with people at all levels of government – federal, state, county and city – both within the US and in the rest of the world. I've developed tremendous respect and admiration for the highly skilled and dedicated workforce that brings us the information that drives our economy. They know what needs to be done to make change happen. The City Council's focus on data and transparency offers a tremendous opportunity to effect change if well designed and well implemented.

My name is Julia Lane, and I am a Professor at New York University's Robert F. Wagner Graduate School of Public Service. I have worked with government data for my entire career. I cofounded the US Census Bureau's Longitudinal Employer Household Dynamics Program, based on state level unemployment insurance wage record data, which is the first, and still the only US statistical program established by a researcher. I founded the first remote access secure data enclave for government data at the National Opinion Research Data Center at the University of Chicago. I also established the NYU Coleridge Initiative, which created the Administrative Data Research Facility as well as the Applied Data Analytics training programs. I am an elected fellow of the American Association for the Advancement of Science, the American Statistical Association, and the International Statistical Institute. I have received the Roger Herriot Award, the Julius Shiskin Award and the Warren E. Miller award for my services to statistics and social science.

An effort to focus on data and transparency could effect change if it ensures that the right technology is put in place to protect the confidentiality of data, that agencies are engaged in the effort to change the way in which they do business, and that their staff are empowered and trained to make use of data with the appropriate methods and tools. I expand on these points in my remarks below.

NYU | WAGNER

JULIA LANE
Professor, Wagner School
Professor at NYU Center for Urban
Science and Progress
NYU Provostial Fellow for Innovation
Analytics

Robert F. Wagner Graduate School of
Public Service
New York University
295 Lafayette Street, 2nd Floor
New York, NY 10012

P: 646 208 4990

julia.lane@nyu.edu

First, the technology exists to enable agencies to share data securely and responsibly. Investments have been made at the Federal level to support the decision making of the Commission on Evidence-Based Policymaking. Ten of the Commission's recommendations have been signed into law: H.R. 4174, the "Foundations for Evidence-Based Policymaking Act of 2018," which improves evidence-based policy through strengthening Federal agency evaluation capacity; furthering interagency data sharing and open data efforts; and improving access to data for statistical purposes while protecting confidential information. Much has been learned from important national investments in the area of Secure and Trustworthy Cyberspace as well as the National Privacy Research Strategy. A great deal has been learned aobut risk management, ensuring that information flows and use are consistent with privacy rules, and the human aspects of building secure environment. Useful strategies that exist to improve computer and cybersecurity and there has been an important debate about how best to securely release aggregate information, particularly whether new cryptographic techniques, such as differential privacy, can be broadly applied.

The NYU Coleridge Initiative's Administrative Data Research Facility (ADRF), which was established less than three years ago to inform the Commission's decision making, already provides a secure environment within which many agencies from jurisdictions around the country have shared data. My colleague, Julia Stoyanovich, as well as others in the community, has deep expertise in the burgeoning new area of fostering responsible data science, which fosters critical literacy, establishes values and beliefs as an explicit part of systems design and implementation, advances new technical research in methods, systems, and algorithms that reflect societal values and relevant laws, and reimagines new systems in data management and data sharing, which incorporates sociological questions about the people affected by algorithmic decision systems. The City Council could provide a list of approved secure environments within which agencies could share data.

Second, shared data is the necessary first step to better transparency, but it is not sufficient to effect change. Because most citizens encounter multiple city and state programs, it is necessary to link data across agency lines to fully understand the dynamic interaction of government with citizens. Agency capacity must be built to understand the many challenges associated with linking data cross agency lines, and to share knowledge about how best to address those challenges. Too few government employees have the requisite skills, and governments often do not have the salary flexibility to compete with the private sector to hire and retain enough in-house data analyst. Confidentiality rules often limit datasharing. Absent a clearly sufficient value proposition, it is difficult to obtain the resources necessary to surmount the many legal and technical hurdles that prevent cross-agency data collaborations. The time to reach agreement can take years—10 years in at least one case!.

These combined challenges have led to the current Catch-22: because they cannot demonstrate the value of new data products, agencies cannot get the significant resources necessary to make use of linked data, but lack of resources mean that they cannot demonstrate value. In human terms, the cost of not combining data (in a timely fashion) is also evident. Dr. Leana Wen, Commissioner of Health, City of Baltimore has noted, as "part of Child Fatality Review, department heads in Baltimore City government get together once a month. We review every child death that happened in the city since the previous meeting. We ask what more we might have done to prevent that tragedy. In many cases, each of us has a file on the child or the family at least an inch thick. It's

JULIA LANE
Professor, Wagner School
Professor at NYU Center for Urban
Science and Progress
NYU Provostial Fellow for Innovation
Analytics

Robert F. Wagner Graduate School of
Public Service
New York University
295 Lafayette Street, 2nd Floor
New York, NY 10012

P: 646 208 4990

julia.lane@nyu.edu

tragic to compare notes after the child has died—what more could we have done when the child was alive?"

Our experience is that training classes that bring agency staff together to solve such challenges in the context of addressing a common problem is critical to building such capacity. We have trained over 300 government agency staff from over 100 agencies across the country in executive education style classes. Three projects are highlighted on our website (https://coleridgeinitiative.org/training) and are available for download. The title of each of them show the nature of the work: "From Prosecuted to Job Recruited: An Exploratory and Machine Learning Approach to Employment after Prison"; "Addressing Recidivism: Intervening to Reduce Technical Violations and Improve Outcomes for Ex-Offenders" and "Mommy Don't Go: Predicting and Preventing Recidivism of Mothers in the Illinois Criminal Justice System." The City Council could enhance the capacity of agency staff to work with confidential data by encouraging agencies to send staff to similar executive education programs so that they learn the core data science skills necessary to do their jobs in the new data driven environment that we all face.

Third, much work has also been done to develop aggregate indicators that are understood by the agency and that are robust in construction, in health, housing, transportation and criminal justice, to name just a few. Agencies that are empowered to generate additional indicators, traceable to the underlying source data, will better understand the causes of, and how to respond to, changes in the indicators. Community feedback is, of course, essential. The aggregate data indicators could be disseminated to the community for comment and iterative feedback so that effective evidence-based policy can be developed around sensible aggregate indicators. The City Council could encourage agencies to be involved in such a feedback loop, and reward them for doing so.

In short, I recommend that City Council leverage substantial existing work that has been done at the federal and local level in three ways.
(1) Ensure the right technology is put in place to protect confidentiality for data sharing.
(2) Provide professional development funds adequate to ensure that agencies are engaged in the effort to change the way in which they do business.
(3) Reward agencies that substantively engage with, and receive input from, the community in developing critical indicators.

NEW YORK UNIVERSITY

Julia Stoyanovich
Assistant Professor of Computer Science and Engineering, Tandon School of Engineering

Assistant Professor of Data Science, Center for Data Science

Department Computer Science and Engineering
Tandon School of Engineering
2 MetroTech Center
Brooklyn, NY 11201

P: 212 998 3251
M: 917 470 8199

stoyanovich@nyu.edu

Testimony of Julia Stoyanovich before New York City Council Committee on Technology and the Commission on Public Information and Communication (COPIC)

February 12, 2019

Dear Speaker Johnson, Chair Koo, and members of the Committee and Commission:

The data revolution is transforming every sector of science and industry, but has been slow to reach local and municipal governments that deliver vital human services in health, housing, and mobility. The opportunities of data-driven algorithmic decision making in urban contexts have long been recognized, evidenced by the remarkable progress around open data, the digitization of government records and processes; and, perhaps most visibly, smart city efforts that emphasize using sensors to optimize city processes. Despite this progress, the public sector is slow to adopt data-driven technology, for two related reasons, both highly relevant to the topic of today's hearing.

The first concerns the legal and technical difficulties inherent in the sharing of sensitive data, both among government agencies and with external entities. The second reason is the government's mandate for responsibility — meaning that any decision made by algorithms will need to be scrutinized by the affected individuals, groups, and the general public. *In my testimony today, I will argue that both barriers to adoption of data-driven technology can be overcome by establishing a robust and flexible data-sharing infrastructure.* Consequently, establishing this infrastructure should be seen as a clear strategic and operational priority for New York City.

My name is Julia Stoyanovich, I am a resident of New York City (District 7). I hold a Ph.D. in Computer Science from Columbia University. I am an Assistant Professor of Computer Science and Engineering at New York University's Tandon School of Engineering, and an Assistant Professor of Data Science at the Center for Data Science. In my research and teaching, I focus on *responsible data science* — on incorporating legal requirements and ethical norms, including fairness, accountability, transparency, and data protection, into data-driven algorithmic decision making.[1] Some of the students enrolled in my Responsible Data Science course are here today.[2]

I am an appointed member of a Task Force established in response to Local Law 49 of 2018, in relation to automated decision systems used by agencies (the ADS Task Force). Opinions in this testimony, while informed to some extent by my work on the ADS Task Force, are my own, and do not represent the views of the Task Force.

I am thrilled that New York City is maintaining its leadership role in responsible data-driven governance. We are the first, and only, US city to pass an Open Data Law (Local Law 11 of 2012) and an Automated Decision Systems Law (Local Law 49 of 2018). Further, Local Laws 245 and 247 of 2017 establish the role of the Chief Privacy Officer, in support of responsible citywide data sharing

---

[1] See https://dataresponsibly.github.io/ for information about this work, funded by the National Science Foundation through the BIGDATA program (NSF Award #1741047).
[2] DS-GA 3001.009 Responsible Data Science, all course materials are publicly available at https://dataresponsibly.github.io/courses/spring19/

practices. Committee on Technology and the Commission on Public Information and Communication have an imperative to act jointly, to continue creating an environment in which legislative efforts and technological innovation act in concert, with the goal of *"improving government transparency, improving the public's access to government information, protecting personal information privacy, and facilitating data sharing between city agencies"* — the topic of today's hearing.

In my statement today, I would like to make three points:
1. Establishing a robust and flexible data-sharing infrastructure should benefit multiple stakeholders.
2. There is a continuum of data sharing modalities between open data and a secure data clean room that need to be explored as part of infrastructure design.
3. Developing a data-sharing infrastructure will require technological innovation, buy-in from city stakeholders, and public engagement.

I now briefly discuss each of these points in turn, and conclude with a set of recommendations. My testimony will be complemented by statements from my distinguished colleagues Julia Lane, Professor at the Wagner Graduate School of Public Service at New York University, and Stefaan Verhulst, Co-Founder and Chief or Research and Development at GovLab, an action research center at New York University.

My first point relates to the importance of establishing a data-sharing infrastructure that benefits multiple stakeholders. *Government agencies* need to share data to make decisions more effectively, and to enact policy in coordination. *Regulators* need access to agency data for purposes of oversight. In both cases, much of the data is sensitive and so is legally encumbered: This data either contains personally identifiable information, or is anonymized but still does not guarantee privacy when linked with other data.[3] Equally as importantly, but discussed less often, is *the public's* need to access data in support of algorithmic transparency.

Recent reports on data-driven decision-making underscore that fairness and equitable treatment of individuals and groups is difficult to achieve[4], and that transparency and accountability of these processes in government are indispensable but rarely enacted[5]. As a society, we cannot afford the status quo: algorithmic bias in administrative processes limits access to resources for those who need these resources most, and amplifies the effects of systemic historical discrimination. Lack of transparency and accountability threatens the democratic process itself.

New York City's ADS Transparency Law (Local Law 49 of 2018) initiates a meaningful response to these threats, and other US municipalities are likely to follow with similar legal frameworks or recommendations. Of utmost importance as this happens is recognizing the central role of data transparency in any algorithmic transparency framework. *Meaningful transparency of algorithmic processes cannot be achieved without transparency of data!*[6] Data transparency in turn cannot be achieved without a robust and flexible data-sharing infrastructure.

---

[3] See https://arstechnica.com/tech-policy/2009/09/your-secrets-live-online-in-databases-of-ruin/ for a descritpion of a now-classic 1997 re-identification attack, in which Latanya Sweeney, a graduate student at the time, re-identified Massachusetts Governor Weld by linking anonymized hospital visit records and public voter rolls.
[4] MetroLab Network, "First, Do No Harm: Ethical Guidelines for Applying Predictive Tools within Human Services," 2017, https://metrolabnetwork.org/data-science-and-human-services-lab/
[5] Robert Brauneis and Ellen P. Goodman, "Algorithmic Transparency for the Smart City," *Yale Journal of Law & Technology* 20, no. 103 (2018), http://dx.doi.org/10.2139/ssrn.3012499
[6] Julia Stoyanovich and Bill Howe, "Follow the Data! Algorithmic Transparency Starts with Data Transparency," *The Ethical Machine*, November 27, 2018, https://ai.shorensteincenter.org/ideas/2018/11/26/follow-the-data-algorithmic-transparency-starts-with-data-transparency

What is data transparency? In applications involving predictive analytics, data is used to customize generic algorithms for specific situations—that is to say that algorithms are *trained* using data. The same algorithm may exhibit radically different behavior—make different predictions; make a different number of mistakes and even different kinds of mistakes—when trained on two different datasets. In other words, without access to the training data, it is impossible to know how an algorithm would actually behave. Decision-making applications that do not use machine learning technology, such as explicitly stated decision procedures like the Public Safety Assessment and Decision Making Framework (PSA), or that do not attempt to predict future behavior based on past behavior, such as matchmaking methods used by the Department of Education to assign children to spots in public schools, are still heavily influenced by the properties of the underlying data — they are designed and validated using data. We cannot understand whether these methods work, and what impacts they have on individuals and population groups, if we don't have access to the training and validation datasets. An important role of a data-sharing infrastructure is to support access to these training and validation datasets for the purpose of inter-agency coordination, auditing and oversight, and transparency and accountability to the public.

My second point is that there is continuum of data sharing modalities between open data and secure data sharing environments like clean rooms (also known as secure data enclaves). My colleague Julia Lane will discuss her extensive expertise in developing solutions of this kind for access to administrative data.

Let me continue with my argument about the need for data transparency in support of algorithmic transparency, and observe that, while we require access to the training and validation datasets of a particular automated decision system, these datasets may well be sensitive and so cannot be easily shared or released to the public. That is, data transparency is in tension with the privacy of individuals who are included in the dataset. In light of this, a data-sharing infrastructure can offer an alternative data sharing modality. When raw datasets cannot be exchanged or released, relevant statistical properties of the datasets can be exposed through statistically similar synthetic datasets or data summaries. These can in turn be generated using state-of-the-art methods to preserve the privacy of individuals included in the data.[7]

In addition to privacy-preserving data sharing techniques, appropriate for environments in which a trusted relationship between stakeholders cannot be established, it is possible to develop access control and usage control mechanisms for trusted environments. A carefully designed data-sharing infrastructure can be made to support multiple such modalities.

My third and final point is brief. When developing a data-sharing infrastructure, we must considering legal, societal, and technical aspects of the challenge. A solution will entail engaging technology experts, building competencies and incentives within the City, and developing governance structures. My colleagues Julia Lane and Stefaan Verhulst will discuss these aspects in their statements.

To conclude, I recommend that the City consider the development of a data-sharing infrastructure as a strategic and operational priority, with the goals of (1) increasing efficiency of delivery of human services, and (2) supporting transparency and accountability to the public, thus increasing the public's trust in government. Developing this infrastructure will require significant investment, which should be amortized so as to benefit multiple City and external stakeholders. Different data sharing scenarios will require different sharing modalities, including open data, privacy-preserving synthetic data and summaries, access and usage control mechanisms, and secure data clean rooms.

---

[7] Haoyue Ping, Julia Stoyanovich, and Bill Howe, "DataSynthesizer: Privacy-Preserving Synthetic Datasets," in *Proceedings of the 29th International Conference on Scientific and Statistical Database Management*, Chicago, Illinois, June 27–29, 2017, 42:1–42:5. https://dataresponsibly.github.io/tools/

⚗ GOVLAB

Leveraging and Sharing Data for Urban Flourishing

Testimony by
Stefaan G. Verhulst,
The GovLab, New York University

before New York City Council Committee on Technology and the Commission on Public Information and Communication (COPIC)

February 12, 2019

Dear Speaker Johnson, Chairperson Koo, and members of the Committee and Commission:

We live in challenging times. From climate change to economic inequality, the difficulties confronting New York City, its citizens, and decision-makers are unprecedented in their variety, and also in their complexity, and urgency. Our standard policy toolkit increasingly seems stale and ineffective. Existing governance institutions and mechanisms seem outdated and distrusted by large sections of the population.

To tackle today's problems we need not only new solutions but also new methods for arriving at solutions.[1] Data can play a central role in this task. Access to and the use of data in a trusted and responsible manner is central to meeting the challenges we face and enabling public innovation.

This hearing, called by the Technology Committee and the Commission on Public Information and Communication, is therefore timely and very important. It is my firm belief that rapid progress on developing an effective data sharing framework is among the most important steps our New York City leaders can take to tackle the myriad of 21st challenges.

My name is Stefaan G. Verhulst and I have been a NYC resident for the last 20 years. I am also the Co-Founder and Chief of Research and Development of the The GovLab, an action research center based at the Tandon School of Engineering, New York University (NYU). Our mission is to improve people's lives by changing and updating governance with new technologies. I am also the lead of The GovLab's Data Program and am delighted to share some of the insights we have gained through our work with a variety of (city) partners on open data and data collaboratives.

I am joined today by some of my distinguished NYU colleagues, Prof. Julia Lane and Prof. Julia Stoyanovich, who have worked extensively on the technical and privacy challenges associated with data sharing. I will therefore avoid duplicating our testimonies and won't focus on issues of privacy, trust and how to establish a responsible data sharing infrastructure, while these are

---

[1] Verhulst, Stefaan. 2017. "From #Resistance to #Reimagining Governance: 6 Shifts That Can Improve the Way We Solve Public Problems." OpenDemocracy. December 12, 2017. https://www.opendemocracy.net/stefaan-g-verhulst/from-resistance-to-reimagining-governance-6-shifts-that-can-improve-way-we-solve-.

central considerations for the type of data-driven approaches I will discuss. I am, of course, happy to elaborate on these topics during the question and answer session.

Instead, I want to focus on four core issues associated with data collaboration. I phrase these issues as answers to four questions. For each of these questions, I also provide a set of recommended actions that this Committee could consider undertaking or studying.

The four core questions are:
- First, why should NYC care about data and data sharing?
- Second, if you build a data-sharing framework, will they come?
- Third, how can we best engage the private sector when it comes to sharing and using their data?
- And fourth, is technology is the main (or best) answer?

Let me start with the first question: **Why should New York City care about data and data sharing?**

As a society we increasingly recognize and accept that opening and sharing data is essential for government transparency—the focus of today's hearings. Open data can help shed a powerful light onto hidden corners of governance, ensuring accountability and transparency and, equally important, empowering citizens in the process.

But I am here today in part to tell you that data sharing has another powerful benefit. Our research at the GovLab shows that, when analyzed and used responsibly, data also has the potential to transform how city government works, enabling more agile and legitimate decision-making as well as more targeted and effective service delivery.[2]

Our research suggests that this process happens through at least four pathways:[3]

First, data can transform governance through **improved situational analysis**, which enables more targeted and effective interventions. An increased ability to access and analyze shared data can allow public officials across departments and agencies to better understand, often in real-time, trends in city activity. It can also help officials understand the geographic distribution of various phenomena, such as population flows and new business activity.

Traffic accidents and fatalities, the reduction of which is promoted by this city through the Vision Zero initiative, are two issues that could be better understood through real-time, anonymized traffic pattern and pedestrian data.[4]

---

[2] Verhulst, Stefaan, and Andrew Young. 2016. "Home Page." Open Data's Impact. Accessed February 2019. http://odimpact.org.

[3] Verhulst, Stefaan, and Andrew Young. 2017. "The Potential of Social Media — Intelligence to Improve People's Lives: Social Media Data for Good." New York, NY: The GovLab. http://datacollaboratives.org/social-media.html.

[4] "Vision Zero." n.d. The Official Website of the City of New York. Accessed February 11, 2019. https://www1.nyc.gov/site/visionzero/index.page.

Second, data collaboration can also provide **better insights into cause and effect of phenomena**, allowing policy-makers to focus on root causes of public challenges rather than their symptoms. For example, one could use street-level images and pedestrian analytics to understand how infrastructure and design choices impact those with reduced mobility.[5] Alternatively, one could look at vehicle movement patterns to recognize and measure the seasonality of different industries across the city.

Data collaboration can also **improve predictive capabilities**, enabling better planning and preparation. Consider, for instance, how emergency management officials might use data on fixed-street asset locations and commercial space usage. Before a large storm, this data might identify which areas are most at risk, either because they lack defenses from the elements or are near businesses that present unique hazards to residents. This information could, in turn, allow officials to use resources to mitigate those circumstances. It might also be used to support adequate planning and preparation, providing first responders with targeted information that could help them save lives if a catastrophe were to strike.

And, finally, data collaboration can improve governance by giving officials better ways to **assess the impact** of government programs and initiatives, thus enabling experimentation and evidence based policy-making. For example, more data on commercial space usage and vehicle movements could help the city better understand the effect of construction on local businesses, allowing officials to offer aid and support where needed. During subway work on the 2nd Avenue subway, for example, local store-owners said business declined between 25 and 50 percent. Data would allow the City to provide a more specific cost estimate.[6]

Based on the above, our overarching recommendation would be to increase awareness among city officials and employees about the value of data in decision-making and governance. Achieving this awareness, we believe will help drive data sharing and data-driven practices across the city.

Toward this broader goal, I have at least two specific recommendations. They are that this Committee may:

- First, call for the creation of an urban evidence base, comprising illustrative case studies capturing the potential and impact of data collaboration within New York City and additional documentation and guidance that would serve to incentivize and empower city officials to do more with shared data. This knowledge hub would be akin to a portal we

---

[5] Blair-Goldensohn, Sasha. 2018. "New York Has a Great Subway, If You're Not in a Wheelchair." *New York Times*, January 20, 2018, sec. Opinion. https://www.nytimes.com/2017/03/29/opinion/new-york-has-a-great-subway-if-youre-not-in-a-wheelchair.html.

[6] Berger, Joseph. 2010. "Work on 2nd Ave. Tunnel Hurting Stores, Owners Say." *New York Times*, October 4, 2010, sec. N.Y. / Region. https://www.nytimes.com/2010/10/05/nyregion/05second.html.

have created at the GovLab that similarly documents the potential of data sharing around the world;[7]

- Second, consider a directory of data science experts within city government that would complement a directory of data sets (collected or acquired by the City). Making the existing expertise more searchable and discoverable, city officials may be in a better position to seek help in establishing the value propositions outlined above.

The second question I want to ask is: **If you build a data-sharing framework, will they come?**

In recent years, governments have spent increasing amounts of time, resources, and effort to make government data accessible with the broad goal of making government more transparent and improving people's lives. Indeed, the evidence (as revealed for example by our work at the GovLab) shows that this is true.

Yet, despite the irrefutable potential of open data, our work also shows that much of that potential remains untapped. This is, quite simply, the result of a certain market failure: Many (possibly most) open data initiatives are not designed with a focus on matching the *supply* of government data to the (possible) *demand* for it. In other words, much of the data released is released without a clear idea of what challenges it might address and of how it might be useful or used. As a result, it sits untouched and unused; its tremendous potential (not to mention the resources involved in releasing it) wasted.

So, in answer to the question "If you build it, will they come?" our answer is a resounding *not necessarily.*[8] Data and data sharing through APIs and other technical means have very real potential. But I would caution, equally, that if data-sharing initiatives are not demand-driven, there is a very real risk of wasted time and resources and, as a result, a general loss of faith in the very idea of using data to solve public problems.

One of our key recommendations to this Committee is the city needs a much better understanding of data demand within and beyond city government. To avoid the types of market failures I've just talked about, government needs to ensure the data-sharing infrastructure proposed addresses a specific and identifiable issue, has a clear sense of the target audience of users, and—importantly—is actually usable by that audience.

Mindful of all these issues, the GovLab last year released what we call an Open Data Demand Assessment and Segmentation methodology that provides several recommendations to ensure that the release of data is more demand-driven.[9]

---

[7] For more information, please see: Verhulst and Young, *supra* note 2.

[8] The GovLab, in collaboration with Reboot, the Mayor's Office of Data Analytics, and the Mayor's Fund to Advance New York City, hosted a panel on this topic on March 6, 2018. See: GovLab Admin. 2018. "If You Build It, Will They Come? A Case for Demand-Driven Open Data." *The Governance Lab @ NYU* (blog). March 12, 2018. http://thegovlab.org/if-you-build-it-will-they-come-a-case-for-demand-driven-open-data/.

[9] Verhulst, Stefaan, and Andrew Young. 2018. "Toward an Open Data Demand Assessment and Segmentation Methodology." New York, New York: The GovLab. http://www.thegovlab.org/open-data-demand.html.

For the purposes of this Committee, I would recommend to call for an exercise to establish demand by identifying key questions and problems the city faces. And, in particular, focus this exercise on identifying problem areas (such as affordable housing or traffic accidents) where there is a clear case data could provide most value.

This demand can be identified in several ways:

- One way would involve every city agency developing a list of top ten questions that could benefit from bringing currently inaccessible data to bear and might, with such data made available for analysis, positively transform the way those agencies achieve their mission.

- This could, of course, also be achieved through a top 100 questions list across city government or various other iterations. The point is to think through specific exercises to help identify real questions and real areas of demand. I am sure this will make the process of releasing data far more productive and useful.

- Once questions are identified and their importance validated, engage in a data audit - i.e. review what data is necessary to answer these questions, and compare that with the inventory of data collected and/or acquired by the City as to enabling the matching of the data supply to demand. This effort could also inform the determination of who has authorized access to which datasets for what purposes.

The third question I want to ask today is: **How can we best engage the private sector when it comes to sharing and using their data?**

The reason we need to ask this question is because, even though many governments and cities around the world have embraced data and the principles of open data, the truth is that much data that could be relevant for policymaking still resides within the private sector. This is as much true for data itself as it is of data capabilities—the ability to process and analyze data, and to derive relevant insights.

So if cities are to address modern public challenges and improve people's lives, they need to find ways to engage with the private sector—and in particular to gain access to privately held data. At the GovLab, we've spent a significant amount of time working on the notion of data collaboratives, which are an emerging form of public-private collaboration in which corporate-sector data is leveraged to help find innovative solutions to public challenges.[10]

This is not the place to go into great detail on what data collaboratives are, the various ways they are organized and what they can do. We have established a repository of more than 150 examples of data collaboratives.[11] Let me just give you a few examples of city data collaboratives in action, and how they've helped solve public problems:

- In Chicago, data collaboratives helped public agencies, newsrooms, academics, and researchers better understand the local criminal justice system through the sharing of

---

[10] Verhulst, Stefaan, Andrew Young, and Prianka Srinivasan. 2018. "Introduction." Data Collaboratives. 2018. http://datacollaboratives.org/introduction.html.

[11] A full listing of these examples can be found at: "Data Collaboratives Explorer." 2018. Data Collaboratives. 2018. http://datacollaboratives.org/explorer.html.

arrest and investigatory stop data, snapshots of the county jail population, and information on the State's Attorney's cases.[12]

- In Los Angeles, a data collaborative between LinkedIn and the Office of the Mayor of Los Angeles gave the mayor's data team anonymized data on the tech talent in the city. This resource informed the workforce and educational policies of Los Angeles as it sought to enhance its technological talent.[13]
- In Singapore, crowdsourced, user-suggested routes helped improve the utility and responsiveness of Singapore's private bus transportation services.[14]

As for specific recommendations on how the city could consider accessing and using private sector data, I have a few, which I will outline here:

- Commission a survey of privately held data sources covering New York City with the goal of better understanding whether they can help answer questions or solve problems identified through the process outlined earlier;
- Identify and establish relationships with key individuals within corporations who are in charge of data and data sharing. We call such individuals "data stewards" and they are critical actors in any data sharing exercise.[15] A New York network of data stewards could help to build momentum and establish good practices around the use of private-sector data assets that could provide value for the City, while also surfacing key challenges dampening the positive public impact of our current data age.
- Finally, we believe that the notion of data collaboratives holds tremendous potential, and we recommend New York stake out a leadership role in testing and refining approaches for unlocking the public value of private-sector data. While many cities are rapidly adopting new "Smart City" tools, New York has an opportunity to support rigorous research and analysis of operational and governance models for data collaboration and benefit from more robust and evidence-based approaches to data-driven governance.

Moving on now to the fourth and final question I said I wanted to discuss today. That question was: **Is technology the answer?**

Anyone who has worked with technology in large organizations knows, as I just said in a slightly different context, the technology itself is often the least of the problems. Resistance to change

---

[12] "Chicago Data Collaborative." 2018. Data Collaboratives. 2018. http://datacollaboratives.org/cases/chicago-data-collaborative.html. "Main Page." n.d. Chicago Data Collaborative. Accessed February 2019. https://chicagodatacollaborative.org/.

[13] "Investing in a Thriving Tech Workforce Pipeline for LA." 2017. Los Angeles Official Website. 2017. https://data.lacity.org/stories/s/Investing-in-a-thriving-Tech-workforce-pipeline-fo/x899-9vc5/. "LA Tech Talent Pipeline." 2018. Data Collaboratives. 2018. http://datacollaboratives.org/cases/la-tech-talent-pipeline.html.

[14] "Beeline Crowdsourced Bus Service." 2018. Data Collaboratives. 2018. http://datacollaboratives.org/cases/beeline-crowdsourced-bus-service.html.

[15] Young, Andrew. 2018. "About the Data Stewards Network." *Data Stewards Network* (blog). November 26, 2018. https://medium.com/data-stewards-network/about-the-data-stewards-network-1cb9db0c0792.

and transformation (to new insights or new forms of innovation) doesn't usually occur because of technical problems. It arises as a result of entrenched institutional and cultural resistance.[16]

This is true of the private sector, and it is equally true of government. Establishing a responsible data-driven environment in the city will require nothing less than a cultural shift within agencies and all aspects of the government. Sometimes the cultural shift is needed to overcome overt resistance. Often, it's required to deal with risk averse behavior. For example, many city officials recognize the potential of data yet might still avoid sharing out of a fear of future penalties and sanctions for sharing that some might claim to be unauthorized or ill advised. There is a general feeling (which is probably largely true) that they are more likely to get into trouble for *too much* sharing than for *too little*.

So, in answer to the question, Is technology the answer? I would argue technology is necessary, but it's definitely not sufficient. We need to consider the whole issue of data sharing and data-driven governance within a much broader context.

Out of a wide range of possible options, I have three specific recommendations today for incentives and training approaches that could help drive cultural change:

- First, this Committee should consider establishing a set of metrics to evaluate and measure agencies' performance with regard to data sharing. These metrics need to include a number of indicators, including discoverability (e.g., through data directories); timeliness, to ensure insights are fresh and accurate; and usability, including but not limited to data standards and formats;
- Second, to facilitate the necessary cultural shift, the city government should consider integrating data sharing and use into performance reviews and budget and personnel resource allocation for all agencies. This is an important step in creating the right incentives.
- And, finally, consider creating a "data stewards" award that would be given to the agency or individual within an agency most successful at making data discoverable and leveraging that data to improve their mission. This could also include private-sector actors who have partnered with the city to establish data collaboratives. Such an award would add needed public recognition and appreciation for the value of data sharing in city government.

Summary and Conclusion:

These remarks cover what I wanted to talk about today. My testimony contains specific recommendations and steps for further action. If I were to summarize or encapsulate these points into four broader recommendations, recommendations that can later be filled out with detailed actions, they would include the following:

- Raise awareness within city agencies about the potential positive impact of data sharing;

---

[16] Cultural and institutional roadblocks is one of the many considerations we list in our Periodic Table of Open Data's Impact factors. More information on this issue can be found at: Verhulst, Stefaan, and Andrew Young. 2017. "Periodic Table of Open Data's Impact Factors." Open Data's Impact. July 2017. http://odimpact.org/periodic-table.html.
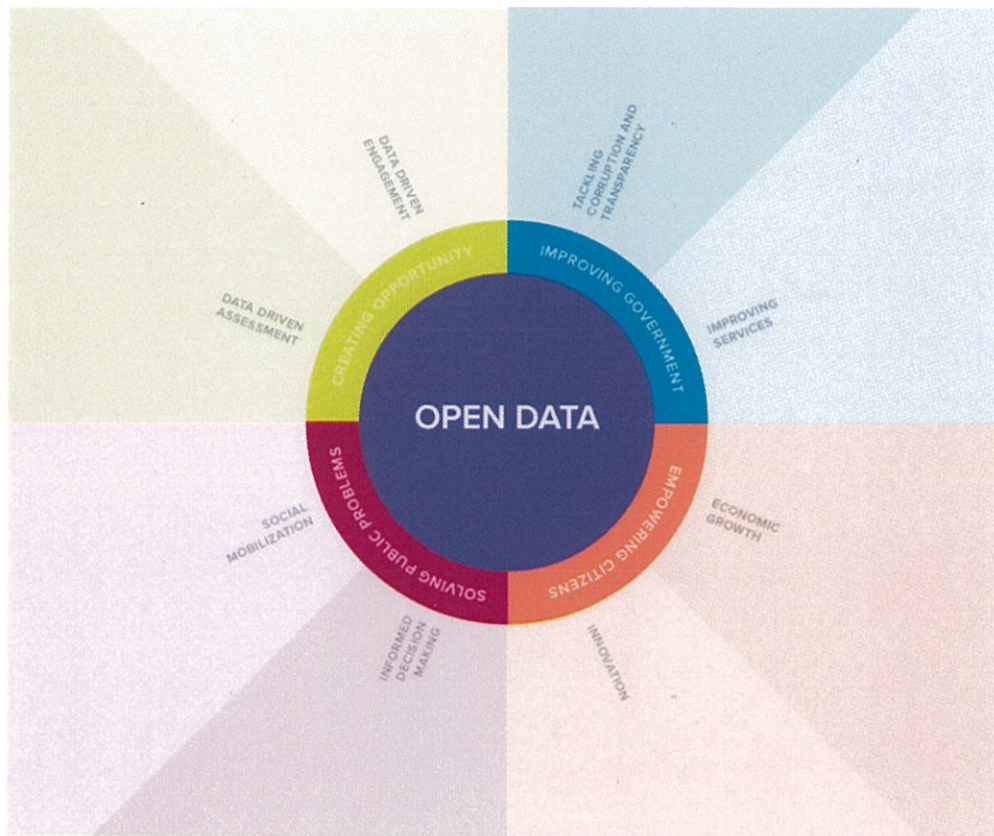
Ā GOVLAB

- Find ways to better match supply and demand in the data equation, and in particular take steps to better understand the demand and need for data;
- Take steps to leverage private data, especially through the use of public-private data collaboratives;
- And finally, facilitate a cultural shift within city agencies to overcome overly reluctant and risk-averse behavior that prevents more data sharing and data-driven decision making.

There are, of course, various pathways to these four recommendations, but the important point is that they are not complicated or financially expensive. The potential of data sharing and collaboration is real and achievable.

What's required to make it happen is leadership—to implement what is already mandated, to overcome bureaucratic inertia, and to transform how decisions are made and services provided. I know that members of this committee are committed to this endeavor, and I look forward to seeing the very positive results on life in this city. In the meantime, I'm happy to answer any questions.

![GOVLAB]

Appendix 1: Taxonomy of Open Data Impact

Based on insights derived from 19 case studies, the GovLab has found that open data projects tend to fit into one of four overarching categories:

- **Improving Government**: Boosting the effectiveness of institutions primarily by tackling corruption and increasing transparency, and enhancing public services and resource allocation;
- **Empowering Citizens**: Allowing citizens to take control of their lives and demand change by enabling more informed decision making and new forms of social mobilization, both in turn facilitated by new ways of communicating and accessing information.
- **Creating Opportunity**: Stimulating citizens and organizations by fostering innovation and promoting economic growth and job creation.
- **Solving Public Problems**: Helping policymakers and citizens address intractable problems with new forms of data-driven assessment and by enabling data-driven engagement.[17]

---

[17] Verhulst, Stefaan, and Andrew Young. 2016. "Open Data Impact When Demand and Supply Meet Key Findings of the Open Data Impact Case Studies." ID 3141474. New York, NY: The GovLab. http://odimpact.org/files/open-data-impact-key-findings.pdf.

Appendix 2: Periodic Table of Open Data's Impact Factors

Problem and Demand Definition	Capacity and Culture		Governance		Partnerships	Risks
U User Research						**Pr** Privacy Concerns
C Causes and Context	**Di** Data Infrastructure		**Od** Open by Default		**Dh** Data Holders	**Ds** Data Security
Rf Refinement	**Pu** Public Infrastructure	**Se** Skills & Expertise	**Fi** Freedom of Information		**I** Intermediaries	**Dm** Poor decision-making due to faulty information
Bg Benefit and Goals	**Lp** Tech Literacy & Internet Penetration	**Fl** Feedback Loops	**M** Performance Metrics	**Dq** Data Quality	**De** Domain Experts	**Pa** Entrenching power asymmetries
Da Data Audit and Inventory	**Rb** Cultural/Institutional Roadblocks	**Rs** Resource Availability & Sustainability	**Rm** Risk Mitigation	**R** Responsiveness	**Co** Collaborators	**Ow** Open-washing

Based on the existing literature and case studies, we have developed a Periodic Table of Open Data Elements detailing the enabling conditions and disabling factors that often determine the impact of open data initiatives.[18] Although the importance of local variation and context is, of course, paramount, current research and practice shows that the elements included in five central issue categories — Problem and Demand Definition, Capacity and Culture, Partnerships, Risks, Governance — are likely to either enable or disrupt the success of open data projects when replicated across countries.

---

[18] Verhulst and Young, *supra* note 16.

Appendix 3: Open Data Demand Assessment and Segmentation Methodology
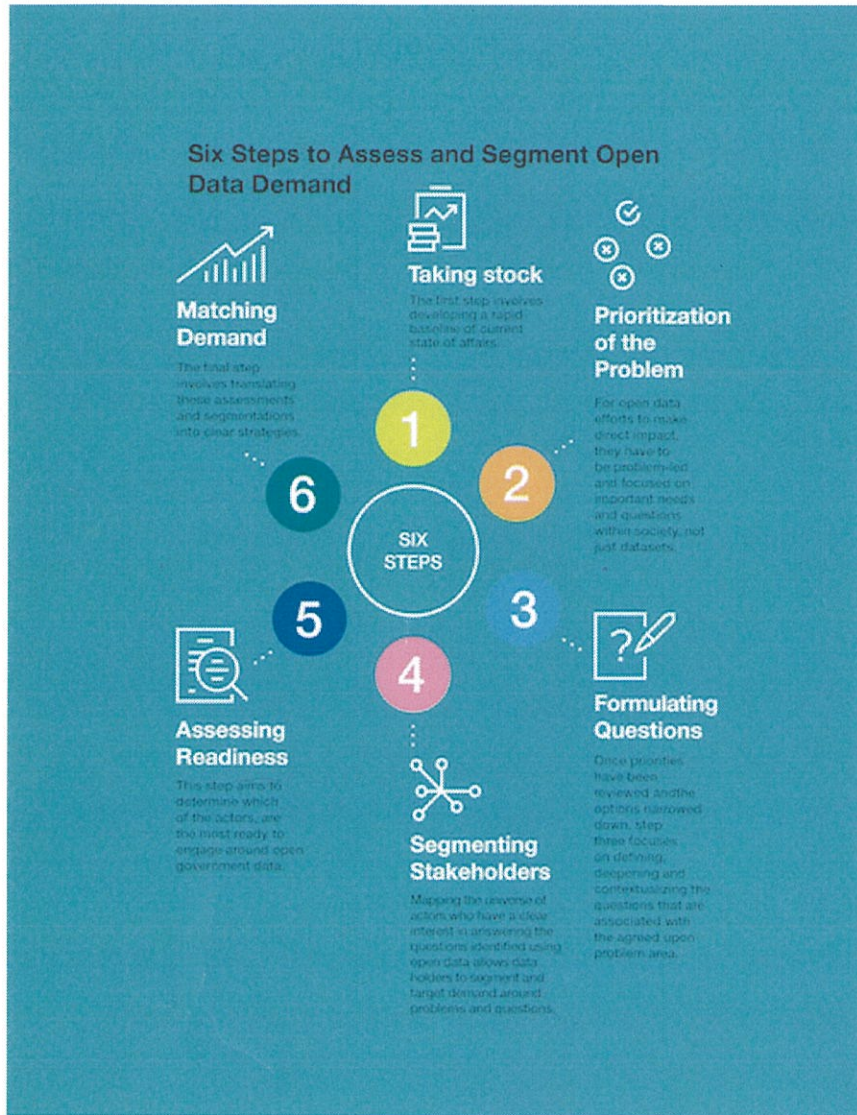
The GovLab, in partnership with the Inter-American Development Bank, and with the support of the French Development Agency developed the Open Data Demand and Assessment Methodology to provide open data policymakers and practitioners with an approach for identifying, segmenting, and engaging with demand. This process specifically seeks to empower data champions within public agencies who want to improve their data's ability to improve people's lives.[19]

---

[19] Verhulst and Young, *supra* note 9.

Appendix 4: Data Collaboratives

DATA COLLABORATIVES HAVE EMERGED AS A NEW FORM OF PUBLIC-PRIVATE PARTNERSHIPS TO ADDRESS SOCIETY'S MOST PRESSING PROBLEMS.

Global Fishing Watch, a partnership between Google, Oceania and Sky Truth, aims to stop illegal fishing by tracking the movement of over 35,000 sea vessels.

Simpa Networks, a company that provides pay-as-you-go solar energy to residents in India, shared its data with DataKind to ensure more people have access to electricity.

NetHope, by acquiring data from the private, public and humanitarian sectors, mapped the trajectory of new Ebola outbreaks in West Africa, **preventing further spread of the virus.**

FOR MORE EXAMPLES, SEE THE GOVLAB'S DATA COLLABORATIVES REPOSITORY datacollaboratives.org/

SIX TYPES OF DATA COLLABORATIVES

Problem: Lack of data diversity leading to ineffective data analysis

Data Cooperatives or Pooling
Corporations and other important dataholders group together to create 'data pools' with shared data resources

Problem: Lack of external actors to apply data analysis skills within public sector

Prizes & Challenges
Corporations make data available to qualified applicants who compete to develop new apps or discover innovative uses for the data

Problem: Limited information and data for academic researchers, stymying their progress

Research Partnerships
Corporations share data with universities and other academic organizations giving researchers access to consumer datasets and other sources of data to analyze social trends

Problem: Inability or lack of resources to create data-driven products to solve a public problem

Intelligence Products
Shared (often aggregated) corporate data is used to build a tool, dashboard, report, app or another technical device to support a public or humanitarian objective.

Problem: Inability to access useful data sets or work with it internal companies like social networks

Application Programming Interfaces (APIs)
APIs allow developers and others to access data for testing, product development, and data analytics

Problem: Lack of expertise to analyze or use private sector data, even when given access

Trusted Intermediary
Corporations share data with a limited number of known partners. Companies generally share data with these entities for data analysis and modeling, as well as other value chain activities.

Data collaboratives are cross-sector collaborations that use data and data science expertise to unlock the societal value of private-sector data.[20] Increasingly, data collaboratives around the world generate insights in health, education, and crisis response, among many other sectors.[21]The above graphics illustrate several notable examples of data collaboratives in the real world and the many forms in which a collaborative can take place.[22]

---

[20] "Data Collaboratives Main Page," *supra* note 12.

[21] "Data Collaboratives Explorer," *supra* note 11.

[22] Verhulst and Young, *supra* note 10.

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____
☒ in favor   ☐ in opposition

Date: _____

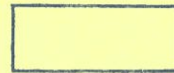**(PLEASE PRINT)**

Name: ANDREW RASIEJ

Address: 31 UNION SQ WEST

I represent: CIVIC HALL .ORG

Address: 118 W 22ND ST

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____
☐ in favor   ☐ in opposition

Date: _____

(PLEASE PRINT)

Name: Albert Fox Cahn, esq.

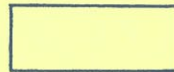Address: _____

I represent: Surveillance tech oversight Proj; STOP

Address: 40 Rector Street

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____
☐ in favor   ☐ in opposition

Date: _____

(PLEASE PRINT)

Name: Julia Lane

Address: 3 Washington Sq Village 10F

I represent: NYU

Address: _____

THE COUNCIL
THE CITY OF NEW YORK

WANTS TO GO TOWARDS END if possible

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor ☐ in opposition

Date: _____

(PLEASE PRINT)

Name: Sumana Harihareswara

Address: PO Box 6542, Astoria, NY 11106

I represent: Changeset Consulting

Address: _____

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor ☐ in opposition

Date: 2/12/2019

(PLEASE PRINT)

Name: STEFAAN VERHULST

Address: 260 18th Street #2 Bklyn 1125

I represent: The GovLAB   Stefaan Verhulst

Address: _____

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor ☐ in opposition

Date: 2/12/2019

(PLEASE PRINT)

Name: Julia STOYANOVICH

Address: 501 W 123 ST, #6D NYC 10027

I represent: NYU

Address: _____

Please complete this card and return to the Sergeant-at-Arms

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor   ☐ in opposition

Date: _2/12/19_

(PLEASE PRINT)

Name: _Gale Brewer_

Address: _Council Chambers, City Hall_

I represent: _Manhattan Borough President_

Address: _1 Centre Street_

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor   ☐ in opposition

Date: _____

(PLEASE PRINT)

Name: _JANET CHOI_

Address: _____

I represent: _MOME/NYC MEDIA_

Address: _____

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____

☒ in favor   ☐ in opposition

Date: _____

(PLEASE PRINT)

Name: _NOEL HIDALGO_

Address: _85 DRIGGS AVE, 11222_

I represent: _BETA NYC_

Address: _____

Please complete this card and return to the Sergeant-at-Arms

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor   ☐ in opposition

Date: _____

(PLEASE PRINT)

Name: David Siffent

Address: 111 Third Ave, NY, NY 10003

I represent: Center on Civil Justice at NYU Law

Address: 139 Macdougal St., NY, NY

Please complete this card and return to the Sergeant-at-Arms

THE COUNCIL
THE CITY OF NEW YORK

Appearance Card

I intend to appear and speak on Int. No. _____ Res. No. _____

☐ in favor   ☐ in opposition

Date: 02/12/19

(PLEASE PRINT)

Name: Angel Diaz

Address: 120 Broadway, New York, New York

I represent: Brennan Center for Justice at NYU Schol of Law

Address: 120 Broadway, New York, New York

Please complete this card and return to the Sergeant-at-Arms