

LEARNING DATA SCIENCE THROUGH CIVIC ENGAGEMENT WITH OPEN DATA

GRAHAM DOVE

BACKGROUND

- This is ongoing research investigating the informal learning that takes place around open government data events and training in New York City (NYC)
- I will present initial findings from interviews with participants across a spectrum of engagements with open data in NYC
- Your feedback on the interpretations that emerge will be greatly appreciated

NYC OPEN DATA

- In 2012 NYC enacted an Open Data law that mandated all public data be made available through a single web portal
NYC Open Data
<https://opendata.cityofnewyork.us>
- Popular data sets include: 311 reports of non-emergency issues, e.g. noise or street cleaning; motor vehicle collisions; City payroll; construction permits; taxi and limousine licenses

NYC OPEN DATA

- An ecosystem of formal and informal groups including municipal agencies, non-profit organizations, commercial start-ups, and engaged individuals act as open data intermediaries^{1,2}
- Develop tools to access NYC Open Data
- Provide open data training
- Undertake analysis to support civic activism
- Host open data events, meetups, and art and design exhibitions

COLLABORATING PARTNERSHIPS

BetaNYC

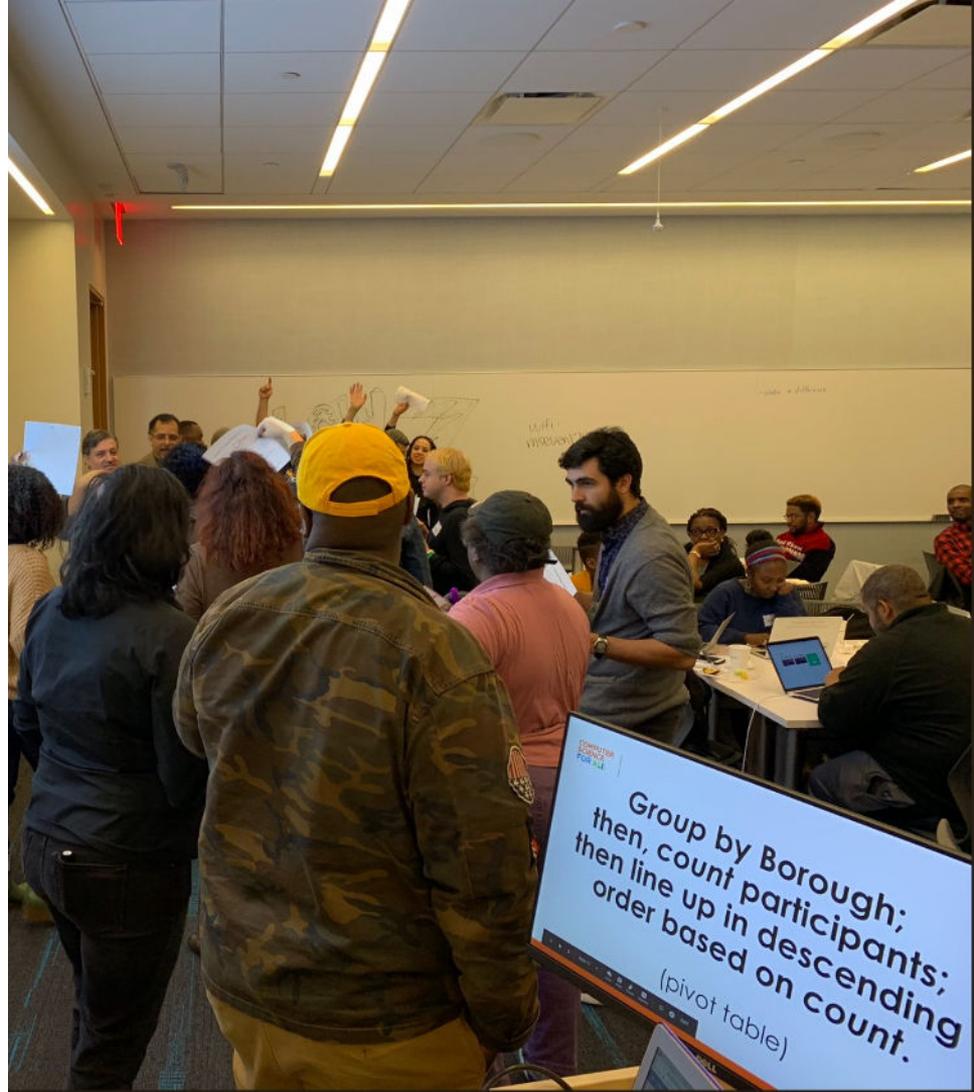
BetaNYC is a non-profit civic tech organization. Founded in 2008 as a meetup group, BetaNYC have developed open data tools, offer open data training, undertake open data research and analyses for community organizations, and organize open data events including NYC School of Data

NYC MODA

The NYC Mayor's Office of Data Analytics (MODA) is one of two municipal agencies responsible for delivering the NYC Open Data program. NYC MODA coordinate open data release and updates, offer training for municipal staff, and organize the NYC Open Data Week festival of events

TRAINING AND EVENTS

BetaNYC and NYC MODA provide open data training events aimed at introducing data analytics concepts to novices through workshops that include activities where participants explore sorting, grouping, filtering, etc. physically by playing the role of data items.



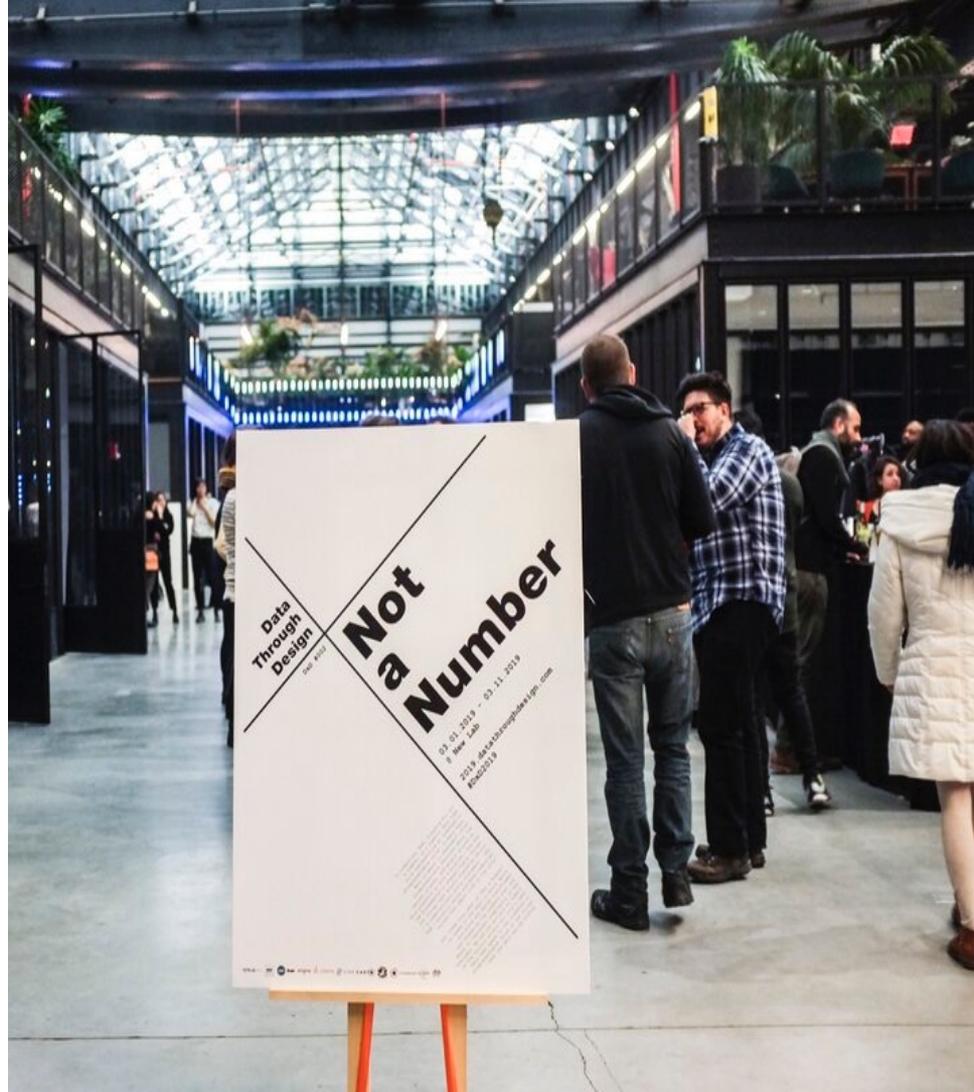
NYC SCHOOL OF DATA

NYC School of Data is a community conference focused on civic tech, open data, and service design for NYC. It provides a forum for sharing and promoting tools, offering training, and developing policy agendas.



NYC OPEN DATA WEEK

NYC Open Data Week is a festival of events that takes place each March to coincide with the Open Data Law anniversary. It takes place across the city and includes discussions, training, and other engagements include the Data Through Design exhibition.



CIVIC INNOVATION FELLOWSHIPS

CUNY Service Corps

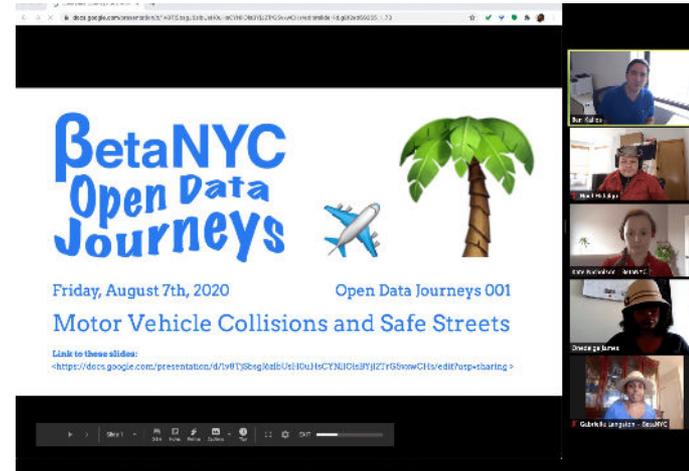
CIF is a program run by BetaNYC and incubated through the Manhattan Borough President's Office in which CUNY Service Corp students learn to use and also teach open data tools and skills, and work with Community Boards to improve their use of data and technology.



VIRTUAL TRAINING AND EVENTS

Open Data Journeys

In response to COVID-19 restrictions on in-person activities, BetaNYC have translated some of their open data training activities into online Open Data Journeys that explore NYC issues through the lens of data. For example, using open data to investigate motor vehicle collisions and safe streets



The screenshot shows a virtual training event interface. The main content area displays the following text and graphics:

- BetaNYC Open Data Journeys** (Logo with a palm tree and an airplane icon)
- Friday, August 7th, 2020**
- Open Data Journeys 001**
- Motor Vehicle Collisions and Safe Streets**
- Link to these slides:**
<https://docs.google.com/presentation/d/7wT3R8qKd6UjH0uHCTNjJc87J2TgShowCHs/edit?usp=sharing>

On the right side, there is a vertical column of five video thumbnails showing participants in a virtual meeting. The bottom of the screen shows a presentation navigation bar with icons for back, forward, search, and other controls.

RESEARCH

- Our aim is to better understand how and why people become engaged with the open data ecosystem in NYC
- What motivates their initial and continued engagement?
- What are the facilitators that can make engagement easier or more pleasurable?
- What are barriers that might prevent or restrict people from accessing and using open data?

DATA COLLECTION

- We conducted semi-structured interviews with seventeen participants with a wide range of experiences of the NYC Open Data ecosystem
- Recruitment included civic tech groups, municipal agencies, community boards, participants from open data training, Civic Innovation Fellows, and attendees of Open Data Week and School of Data events

INTERVIEWS

- We probed participants on how they were introduced to open data and their first experiences
- On their background and experience with regard to data analytics and civics
- On their experiences at events such as NYC Open Data Week and NYC School of Data
- On their perception of barriers and facilitators to accessing and using open data and open data events

MOTIVATIONS

- Open data as a principle of better government to support informed decision making, improve efficiency, and increase civic engagement
- Open data as a resource for public service to help support budget allocation, public health dissemination, and guide policy development

MOTIVATIONS

- Open data as a resource for education, as students practicing data analytics and to support research but also for teaching about real world scenarios
- Open data as a resource for community activism, advocating for data to be made available or providing analytics skills to community groups
- Open data as a resource for entrepreneurship providing data for commercial analyses

BARRIERS

- The range of skills required to use open data effectively include computing and data analytics, understanding civics and how government works, and the ability to frame analysis guiding questions
- Inaccessible language and communication challenges that make it difficult to share exploratory analyses and translate complex processes

BARRIERS

- Negative cognitive experiences such as feeling overwhelmed or out of depth, or reminders of previous bad experiences with math at school
- The availability and quality of data, typically because of privacy or security
- Discovery and awareness, particularly around training and events
- A perceived lack of clarity around open data program objectives and evaluation metrics

FACILITATORS

- Exposure to inspirational examples and role models including tools that show creative applications of open data and people that newcomers relate to
- The community of practice that has emerged around open data, including domain experts, civic activists, designers and data analysts
- Open data events and training, which provide a safe space to ask questions

FACILITATORS

- Making examples personally relevant, particularly during training and when showcasing tools, acts as a hook
- Providing dedicated tools and platforms that simplify interactions and support visualization, especially maps, lowers the entry point.
- Developing an interest in civics and attending or participating in local political and community activities

INTERPRETATION AND TAKEAWAYS

Motivational Alignment

We found that participants' motivations broadly reflected the potential benefits often claimed for open data programs: transparency and accountability, service efficiency, civic participation, and entrepreneurship; but with the addition of education. Whether these benefits are realized remains open to question.

INTERPRETATION AND TAKEAWAYS

Data AND Civics

We found that while the need for data analytics skills and training are necessary for accessing and using municipal open data they are not sufficient.

Understanding civics and government, and engagement with local issues were also highlighted as being necessary.

This aligns with how the open data community are a strong facilitator that provide connections to help bridge the two where necessary.

INTERPRETATION AND TAKEAWAYS

Questions to Actions

The potential to connect to personally relevant topics is a key facilitator and motivation for people to engage with open data. Our findings suggest that to make this engagement meaningful requires an understanding of local issues, which helps in framing the question to ask of the data.

It also requires a level of data analytics skill to find meaningful answers in the data. Finally it requires an understanding of civics and how to interact with government and municipal agencies in order to use the answers found in data as the basis to take action

INTERPRETATION AND TAKEAWAYS

Open Data Literacy

Typically understandings of what it means to be data literate have been based in skills from math and statistics. Where local concerns or issues have been invoked it is typically to provide a context or grounding for learners who may struggle with more abstract concepts.

Our findings suggest that when we think about open data literacy the local and civic knowledge should be considered core aspects of equal importance.

INTERPRETATION AND TAKEAWAYS

The Role of Intermediaries

Our findings strongly suggest that community and collaboration are strong facilitators of continued engagement with open data. Because of this we should acknowledge the role that open data intermediaries can play in creating and maintaining connections. To our knowledge this role has not previously been noted, and should be considered

complementary to the more familiar roles of developer, trainer, and analyst for hire.

INTERPRETATION AND TAKEAWAYS

Pathways to Success

Create entry points for newcomers who may feel initially overwhelmed and intimidated by terminology and jargon. Scaffold the transition from newcomer to experienced open data intermediary e.g. by creating a consistent framework for describing events that reflects their target audience and requirements for technical and civic knowhow

INTERPRETATION AND TAKEAWAYS

Pathways to Success

Provide support for learning open data literacy in a holistic way that incorporates civics and everyday politics simultaneously with data science and technical skills. This will support in skills for framing and asking questions, answering them with data, and then acting on those answers politically. It will also provide a framework for exploring the more complex technical skills in a way that can overcome prior negative experiences around math.

INTERPRETATION AND TAKEAWAYS

Pathways to Success

Develop a train-the-trainers model that can capitalize on the skills, knowledge, and goodwill available within open data and civic tech communities. Such a program would aim to provide: (1) role models who were identified by participants as being a strong facilitator; (2) opportunities for teaching open data literacy, as this was identified as important by civic tech interns; (3) entry points into the open data and civic tech communities and opportunities to deepen engagement.

INTERPRETATION AND TAKEAWAYS

Pathways to Success

Involve community members in discussions when decisions about new data sets to be made open are made. Relating data to everyday life was considered a powerful facilitator and including community members in these decisions can further increase this relevancy and enable people to directly focus on data they are interested in as an entry point that is less overwhelming. This may help to organize open data resources and also make them more accessible.

THANKS

- This work was undertaken with colleagues Jack Shanley, Camillia Matuk and Oded Nov
- It has been greatly helped by BetaNYC and NYC MODA
- The research is funded by National Science Foundation grant #2005890