

Understanding the Users of Open Data

Research Findings

Project Background

The success of the Open Data for All initiative will be measured not by the number of data sets published, but by the number of New Yorkers who use that data in their daily lives. Five years after the passage of NYC's pioneering Open Data Law, the Open Data Portal is still perceived as a tool for—and predominantly used by—a niche community of civic hackers and tech journalists. Yet open data has the potential to help a much wider swath of New Yorkers answer questions and solve problems they confront in their daily lives. Broadening our understanding of the users of open data provides a crucial opportunity to amplify the impact of this initiative.

Intro to Reboot

Reboot is a social impact firm devoted to public sector innovation. We are pioneers in adapting methods and tools from the field of design to serve the public good. Today, we work with foundations, companies, media, governments, and non-profits to bridge the gap between governments and the people they serve.

Project Objectives

- Identify new user segments outside the typical open data community that have the potential to use open data to solve community problems
- Create a clear understanding of what existing and potential users need from the Open Data Team in order to access and use open data to solve problems for themselves and their communities

Key Lines of Inquiry

THEME 1:

How do people access and use information to draw conclusions, make decisions, and solve problems?

THEME 2:

What are people's perceptions of Open Data Team resources, both generally and related to open data?

THEME 3:

What do users and potential users understand about open data?

THEME 4:

What does using open data look like?

THEME 1:

How do people access and use information to draw conclusions, make decisions, and solve problems?

- What factors contribute to people deciding to seek out information to draw conclusions, make informed decisions, and solve problems?
- How do people develop new behaviors around information-seeking related to decision-making?
- How do users' perceptions of open data compare to perceptions of other resources they use to solve problems?

THEME 2:

What are people's perceptions of the Open Data Team's resources?

- How effective do users believe government is at sharing information with the general public?
- How do users view the possible positive/negative impacts of government data initiatives?
- What are common associations with the NYC Open Data brand?
- What additional resources related to open data do people expect the Open Data Team to provide?
- What are perceptions of the NYC open data community?

THEME 3:

What do users and potential users understand about open data?

- What do people associate with open data? Why?
- What do users believe open data is valuable for? What is it not suited for?
- What do they understand about what open data is and how it operates?
- What questions do people have about open data?

THEME 4:

What does using open data look like?

- In what contexts do people use and access open data, and for what needs or reasons?
- How frequently do people access open data? For what duration? How often are they successful?
- What limits people's use of open data?
- How did users become familiar with open data?
- What does successful data use look like? What does it achieve?

Methodology

Landscape Assessment

This research was conducted over a period of five weeks and was built on a desktop review of previously conducted research, and interviews with experts to surface trends and gaps in open data research. Through this research we discovered that much has been done to segment and understand who uses open data portals but little had been done to understand how people were operating within a larger open data ecosystem.

Specific activities included:

- Review of relevant open data research specifically focused on user segmentations related to open data
- Key informant interviews with open data experts:
 - Open Data NYC staff
 - Leaders of the open data community
 - Known “power” downloaders
- Observation of two open data week events

Design Research

The approach is grounded in the discipline of applied ethnography which focuses on developing a deep understanding of stakeholders through semi-structured conversations. By conducting interviews in this style, research teams ask targeted questions that relate to each research theme while also making room for respondents to surface topics and concerns of most interest to them. This allowed for respondent answers to guide the research, rather than our lines of inquiry alone.

- The research team sought to understand how users and potential users make decisions and solve problems most broadly to gain a holistic understanding of these users' operating contexts. Specific methods included:
 - Semi-structured, open-ended conversations, generally one hour in length
 - Observation of how people access and use open data, general data, and information

Targeting Respondents

The research team conducted 20 face-to-face interviews across the five boroughs of New York City. To target a representative sample of users, the research team developed criteria to guide respondent outreach. These criteria were based on initial desktop review and supported by our experience in open government innovation.

Because of the relatively small sample size, the criteria were designed to surface a wide range of respondents with different experiences with open data and varying degrees of community engagement.

Outreach was conducted through key informant interviews, utilizing extensive networks in community advocacy and organizing, cold outreach, and snowball sampling.

TARGETING CRITERIA

- Equal number of respondents with low, medium, and high open data usage:
 - **Low:** has never accessed or heard of the open data portal
 - **Medium:** has browsed the open data portal and downloaded 1-5 data sets
 - **High:** has extensive experience with accessing the open data portal
- Equal number of respondents across four different “problem-solving” profiles:
 - **Community members** organizing around specific issues
 - **Individuals** solving personal or passion problems
 - **Professionals** investigating and exposing issues
 - **Businesses** making decisions to optimize operations

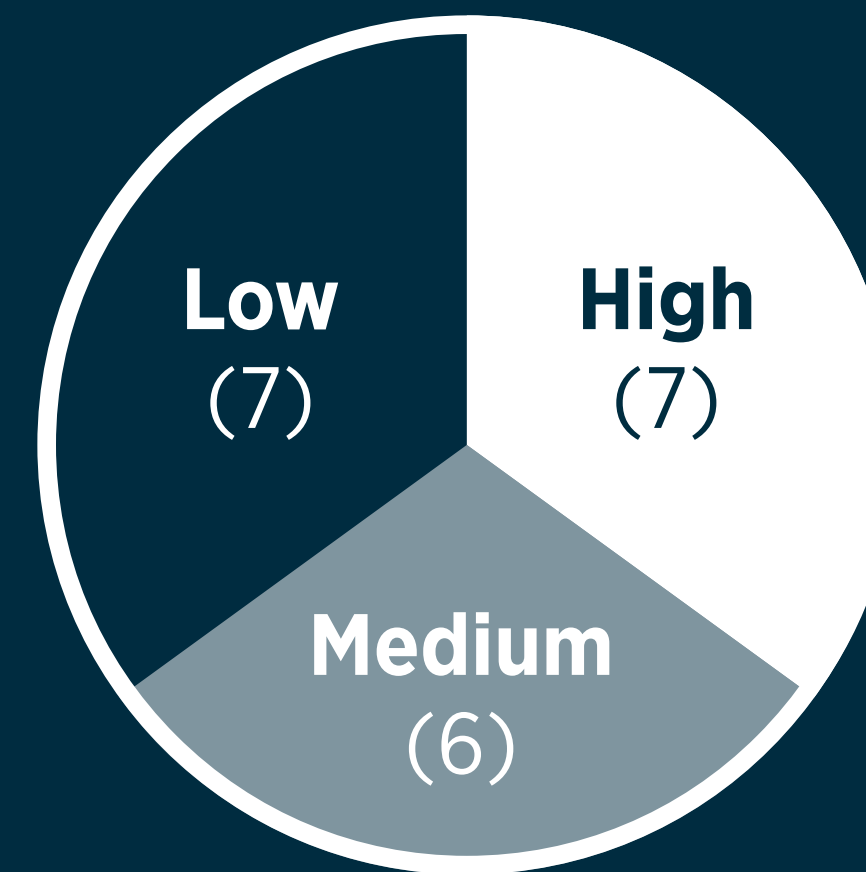
Final Respondent Profiles

Reboot prioritized respondents with diverse and representative backgrounds.

We spoke with:

- Members of the civic tech community including startup owners, platform developers, and urban planners.
- A wide range of individuals within the non-profit sector. Organization specialty spanned:
 - Fine and performing arts, online education, workforce development, participatory budgeting, housing advocacy, and immigrants' rights.
- High school students, educators, and administrators.

DATA SKILL LEVEL:



LOCATION:

5
Boroughs

USER CATEGORY:

Communities Organizing



Individuals Problem-Solving



Businesses Operating



Professionals Investigating



Overarching Findings

Open data...

Creates a buzz

Inspires businesses

Fosters equitable decisionmaking

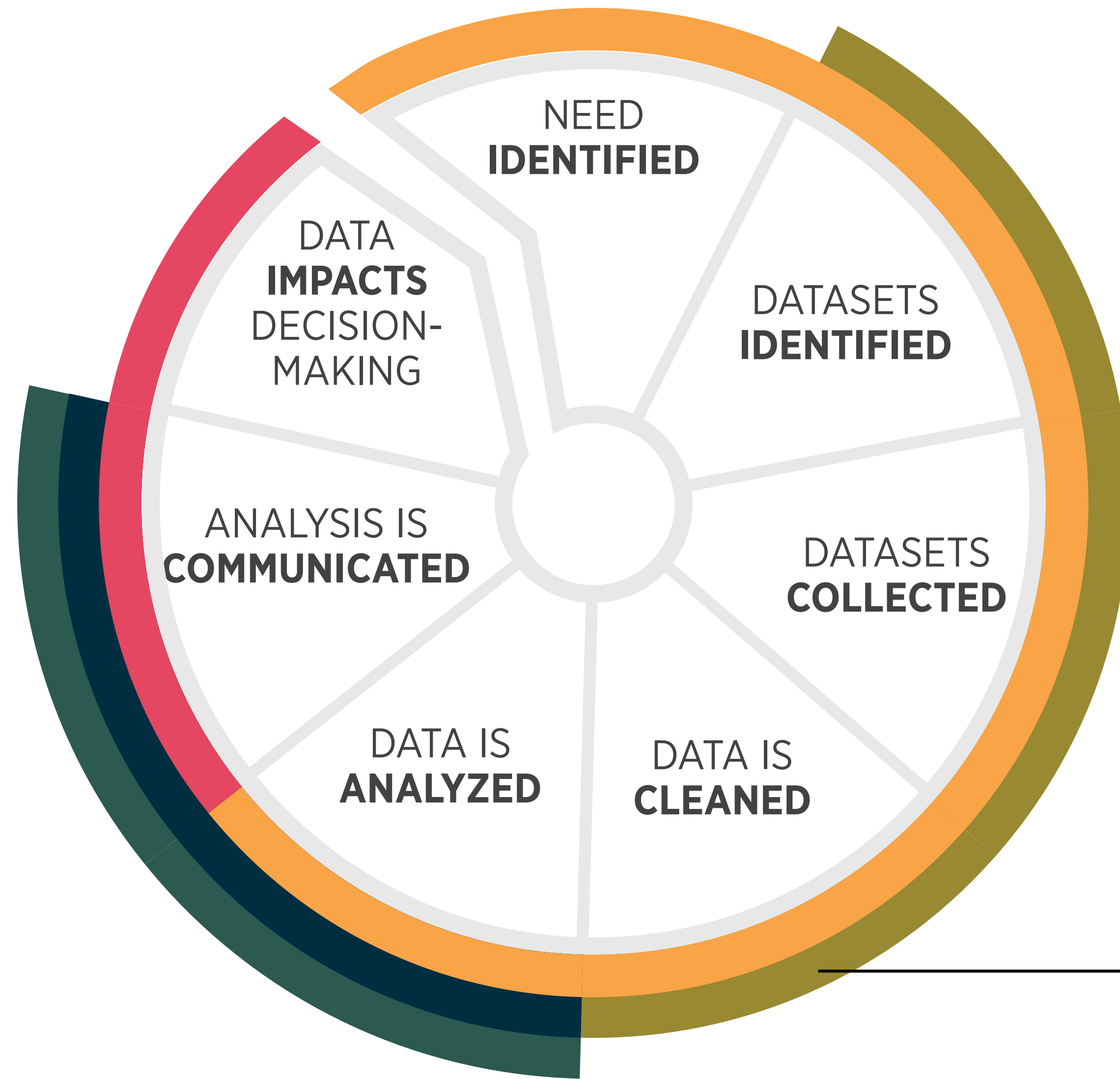
Reveals the makeup of a place

A gateway for learning

Persuades policymakers

Facilitates personal decisions

Open data is a
tool that empowers



The Impact Cycle

Data Skills Required

Technology | **Collection** | **Mapping**
Communication | **Analysis**

Why user personas?

A user persona is a sketch (often with narrative and graphic elements) that represents a composite of multiple people who share common traits and stories. Designers employ personas to conceptualize customers as real people (with needs, motivations, ambitions, and histories) rather than in abstract categories such as “citizens” or “consumers.” But personas are not only helpful for designers.

These tools are intended to help the Open Data Team think in new ways about types of

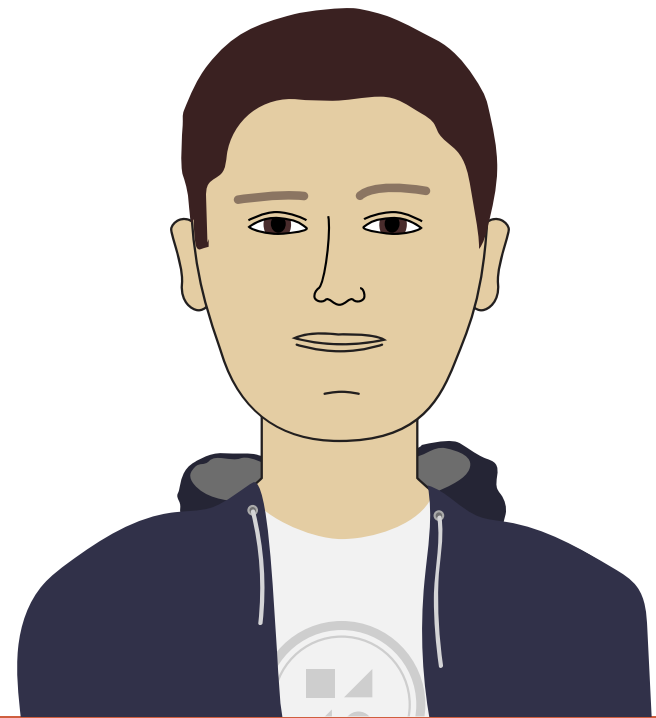
people who use open data. Defining clear pictures of the people who participate at different points of the open data impact cycle will help the Open Data Team design services and resources that will meet the needs of these diverse users.

These profiles are not exhaustive or even mutually exclusive, they simply represent the types of actors who play a role in helping data move from spreadsheet to impact. You may even recognize yourself in one or more of them.

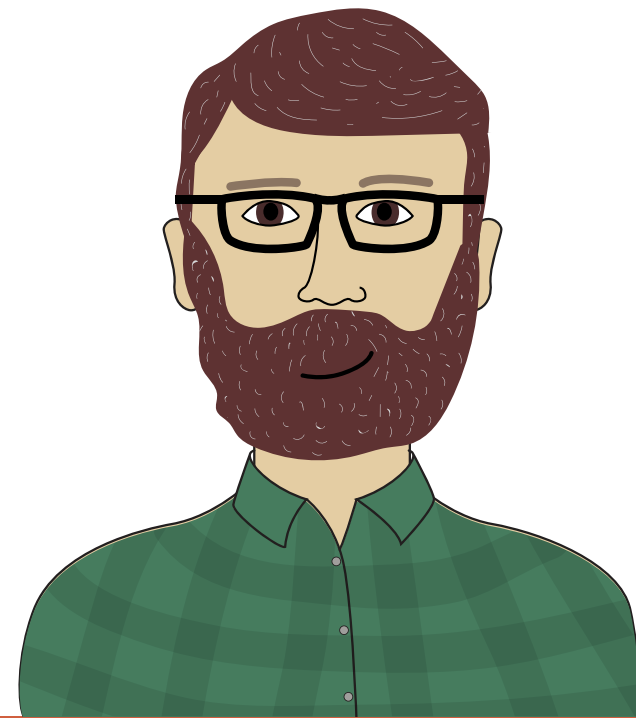
Meet the users of NYC's open data.

CURRENT
OPEN DATA USERS

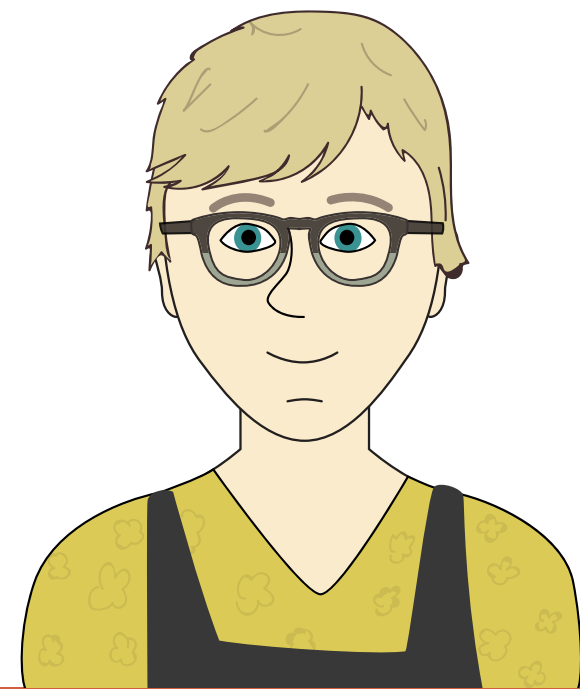
POTENTIAL
OPEN DATA USERS



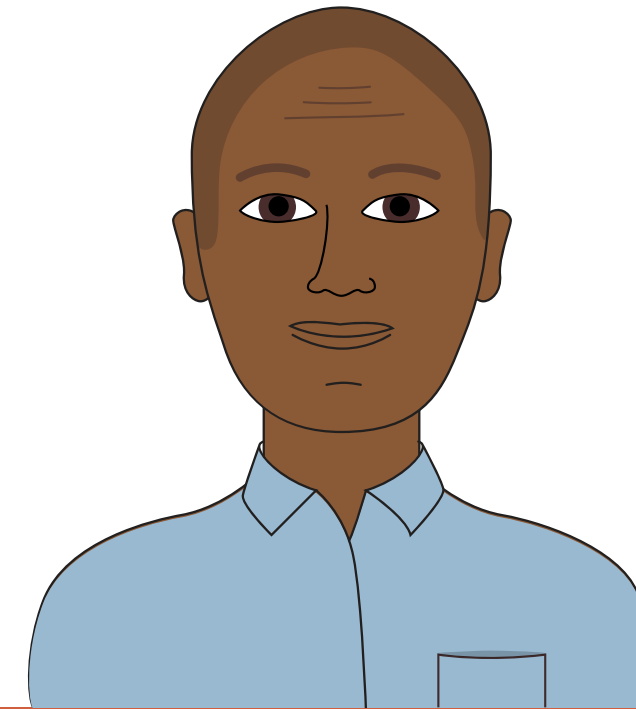
KENT
**Meticulous
Mapper**



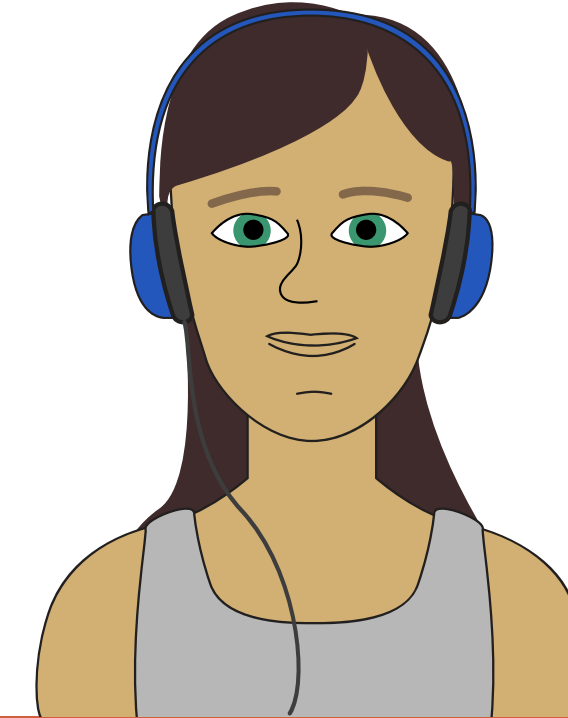
ROMAN
**Local
Liason**



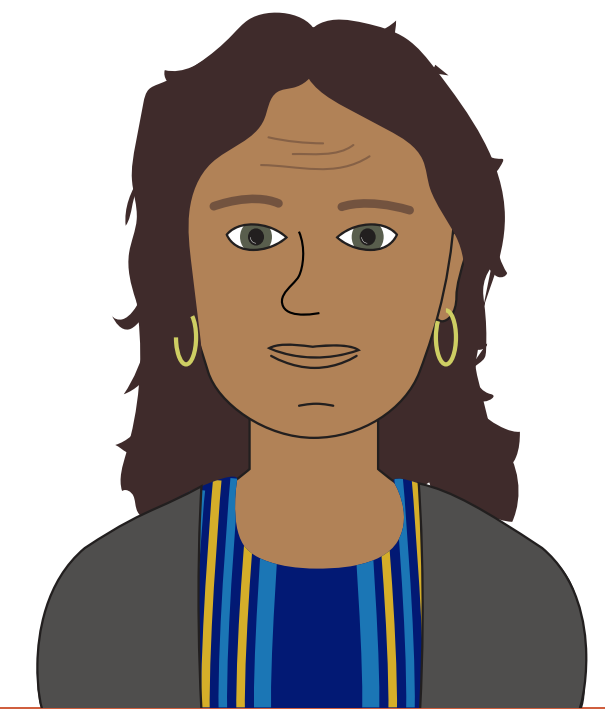
JUNE
**Influential
Interpreter**



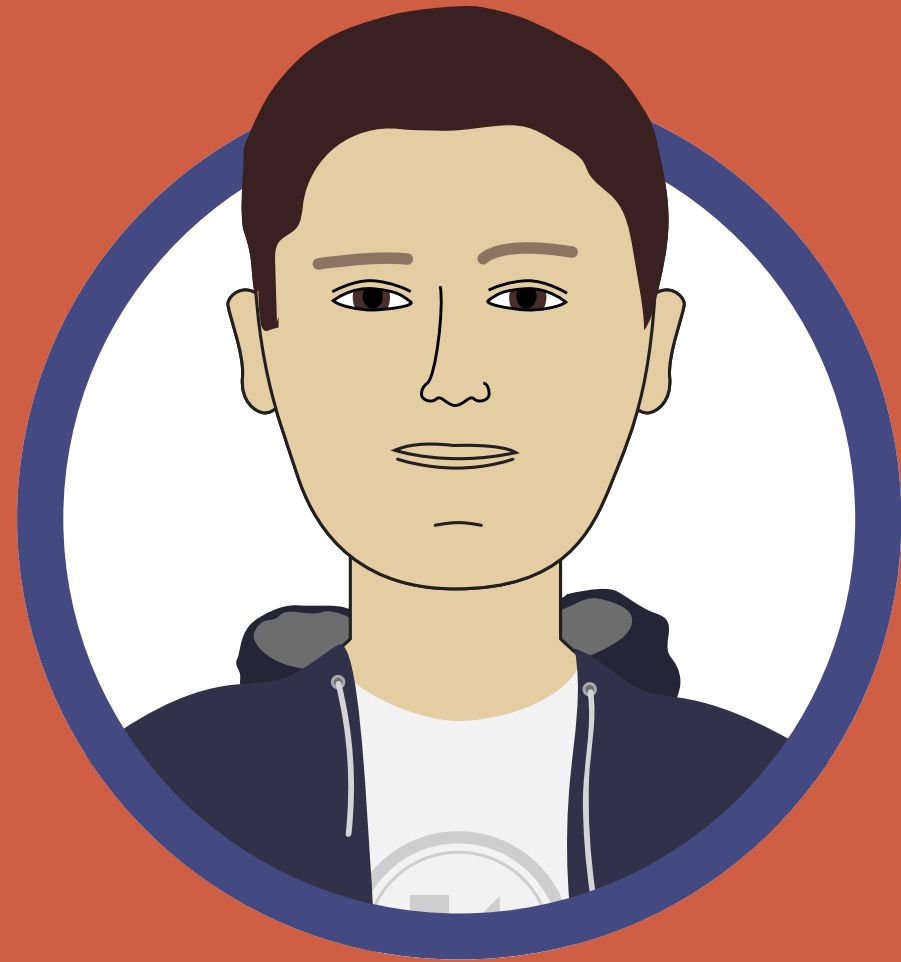
MARCUS
**Equipped
Explorer**



SARAH
**Busy
Bystander**



ALANA
**Community
Champion**



KENT Meticulous Mapper

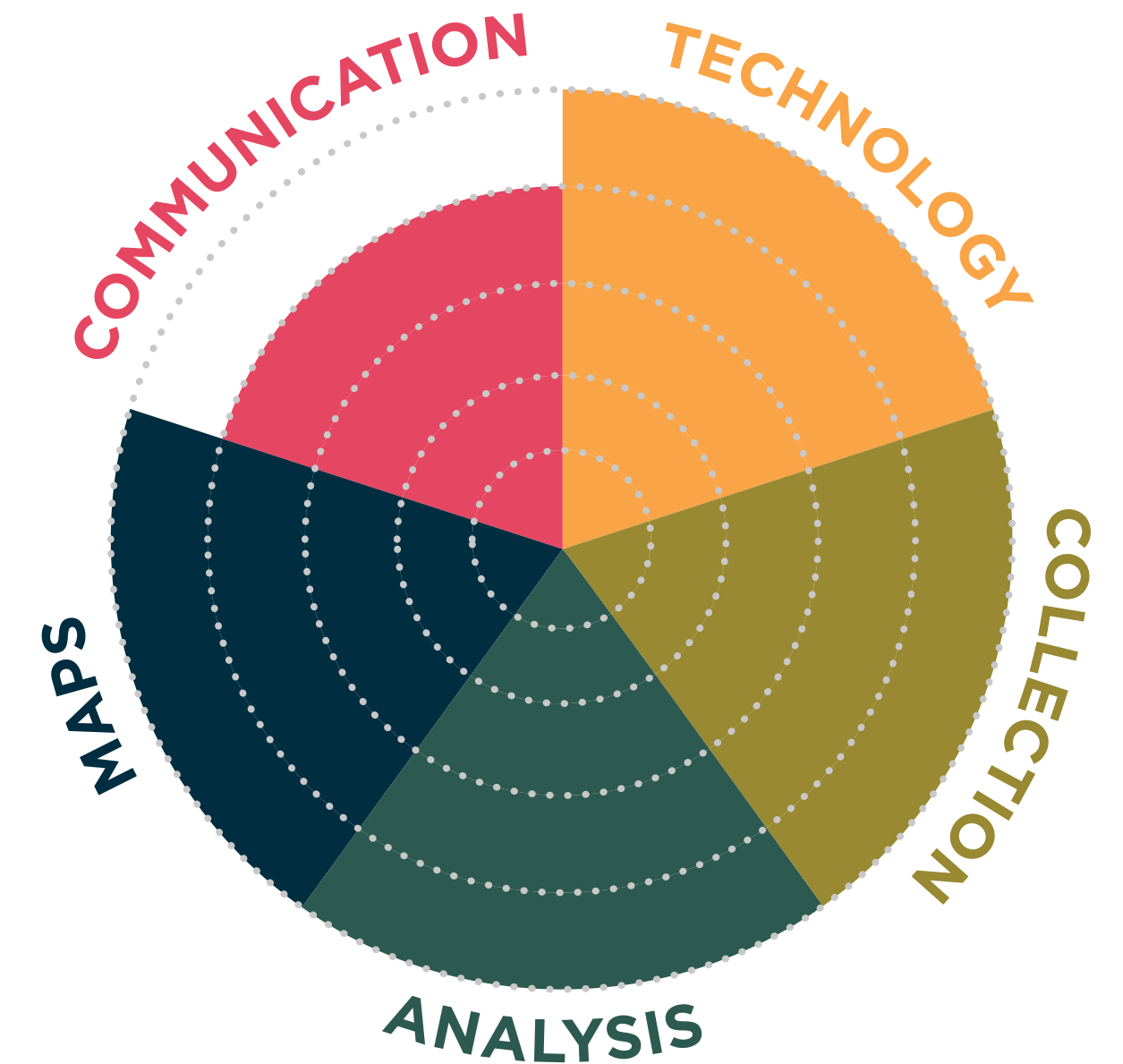
AGE:
34

OCCUPATION:
Lead Developer at a small
civic mapping startup

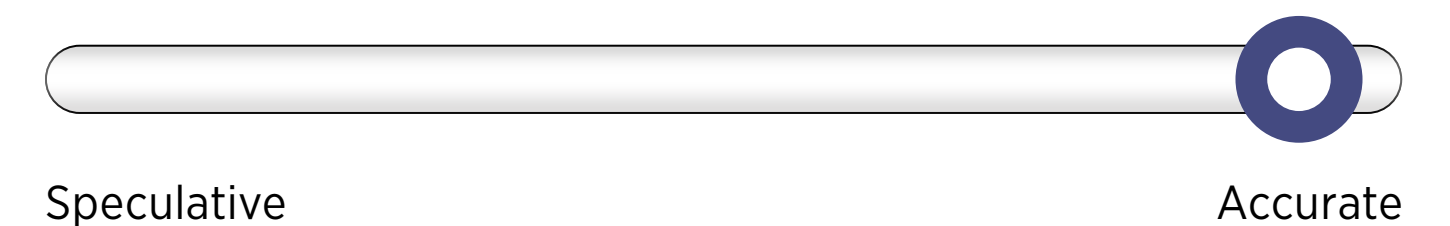
LIKELY LOCATIONS:

At a nearby co-working space, a data and technology MeetUp event, a seminar on GIS at a local university, Civic Hall, a local coffee shop with free wifi, online in his GitHub, or commenting in data-related forums.

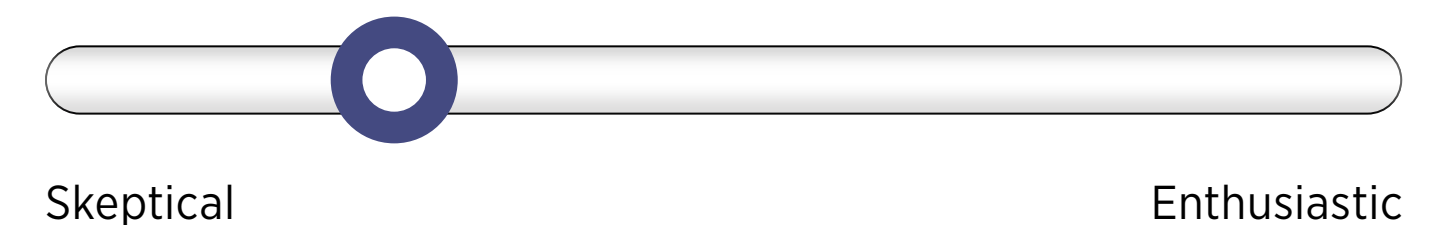
“Open data for all sounds nice, but data doesn’t explain itself. You need interpreters and activists using the data to create stories. Hopefully if people have tools to use the data the stories will come out.”



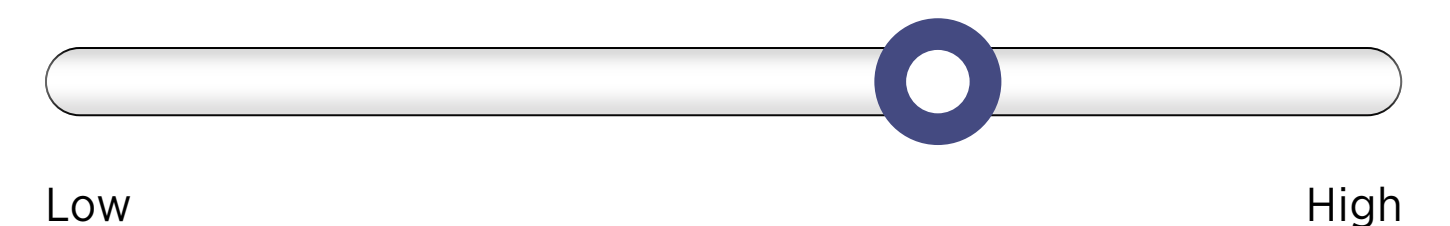
Understanding of Open Data

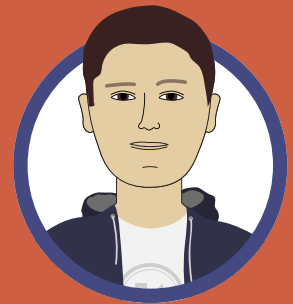


Attitude Towards Open Data



Influence on Community





KENT

Meticulous Mapper

Background:

Kent has been passionate about civic tech since he discovered it while studying Urban Planning at University of Pennsylvania. After taking a class in GIS, he improved his data skills through YouTube videos and online forums—finding a Facebook group of Philadelphia-based technologists especially helpful. After working as an urban planner in Philadelphia for a few years he moved to New York City.

In New York, he attended monthly meetups of people working in civic tech in the city, such as those organized by BetaNYC and Civic Hall. At one of those events he met the founder of the civic mapping startup where he currently works. He values the opportunity to merge his interests in urban planning and civic tech through his job.

Kent stays up to date with events and projects in the local civic tech community through newsletters and Slack teams. He regularly updates and builds out new side projects in Github. Most of the time his projects are inspired by interesting datasets he comes across, but sometimes he responds to requests. Recently, someone from the Department of Education asked him to create an app to map schools with the lowest attendance levels.

Experience with Open Data

Kent is very familiar with open data. To update, maintain, and build out the platform he develops as his day job, he pulls APIs from many sources, both open and private. Most often, he pulls City data from Zillow, PLUTO, and Big Apple Bytes. Kent worked with open datasets from the city before the portal was created, and still prefers to go straight to the source, such as the specific agency websites.

He feels that agencies have greater incentive to update their own agency page, and he isn't sure how often the portal datasets are updated or who handles quality assurance for this centralized platform. When searching for a particularly obscure dataset, he will start with a Google search rather than the data portal.

Because of the education app he recently created, he has made connections at a few agencies. He relies a lot on these personal networks to get the data he needs when he can't find it readily available on city agency websites. When he does go through the open data portal he makes sure he uses an app that his friend made to circumvent the portal's front-end interface and download the data directly.



ROMAN Local Liaison

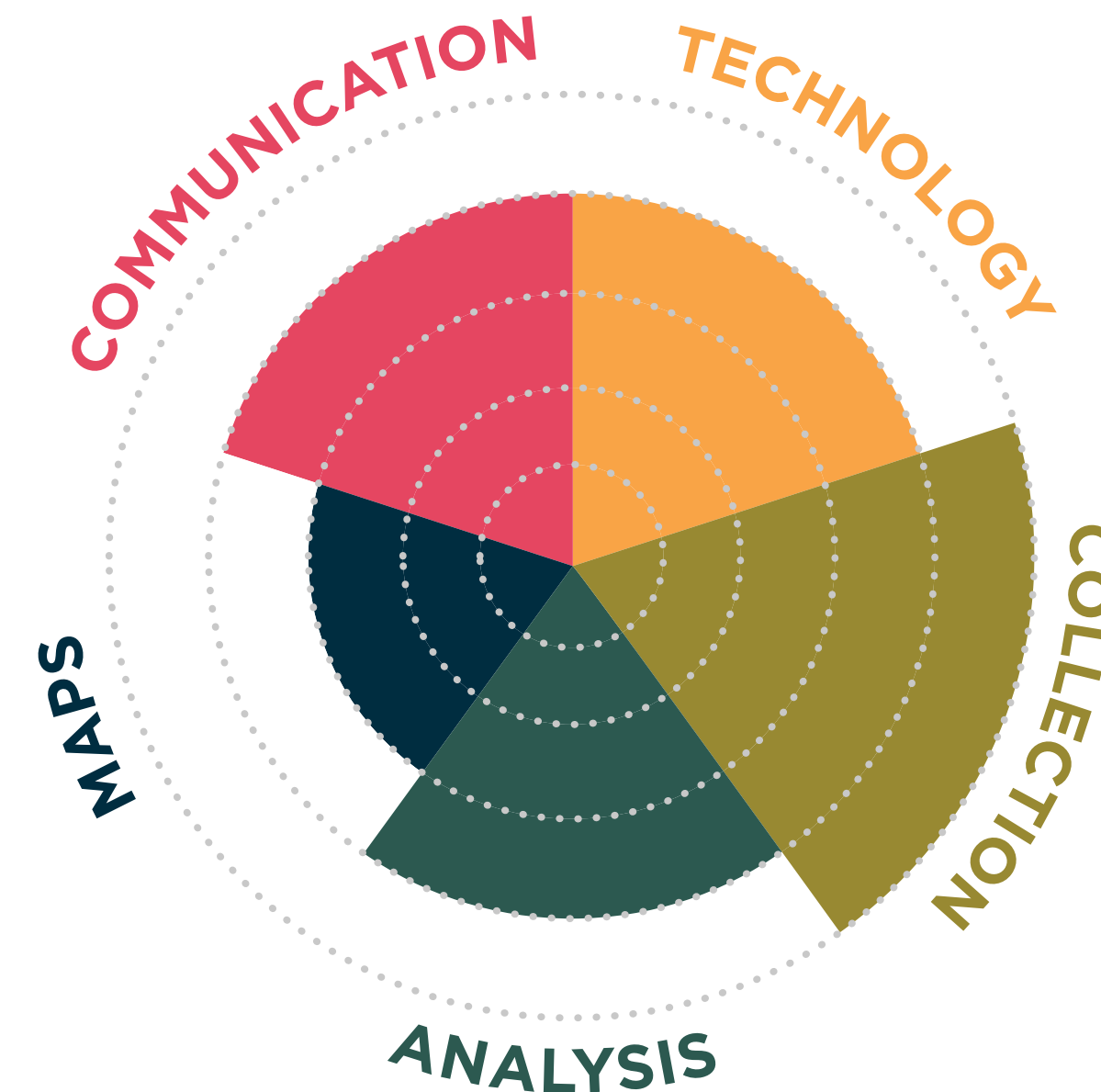
AGE:
38

OCCUPATION:
Project Manager at a
Bronx community
services nonprofit

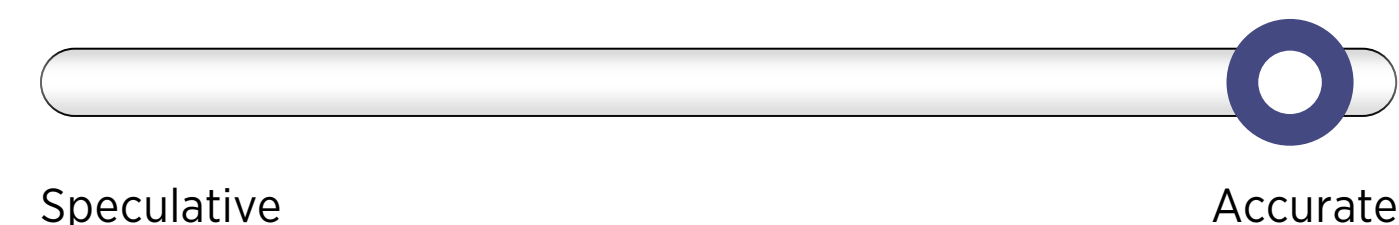
LIKELY LOCATIONS:

Booking a room at his local library, browsing Github, at a MeetUp event focused on a social cause, at a block party thrown by his church, or managing a pop-up table at a community services fair.

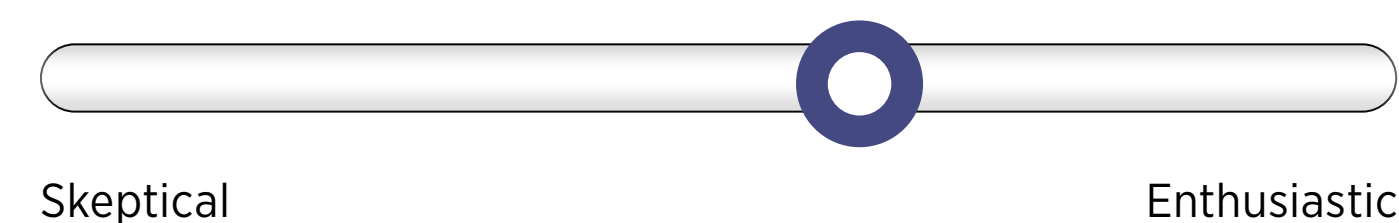
“Open data is in a world full of technical people. I am torn about this, on one hand it's important for the city to provide broad data, on the other hand, it causes average citizens to become reliant on technologists.”



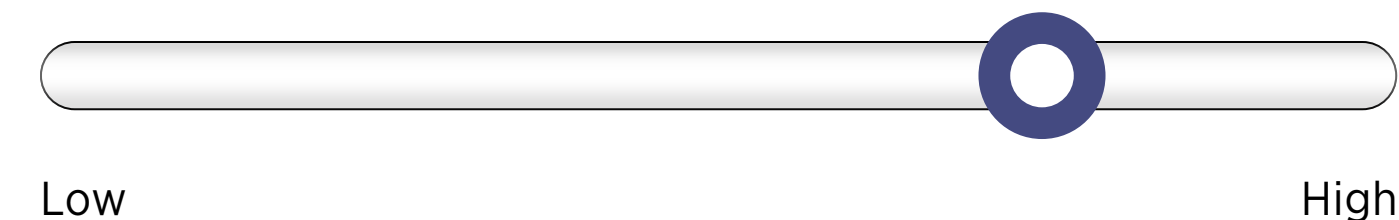
Understanding of Open Data



Attitude Towards Open Data



Influence on Community





ROMAN

Local Liaison

Background:

Growing up, Roman gravitated towards playing with computers and inventing things. Shortly after graduating from college, he began working in data entry for a health services company. He found the job boring, but became interested in tracking his daily entry outcomes and comparing them to his coworkers. This interest in analysis and outcomes led him to work at a nonprofit in his neighborhood.

Roman's problem-solving orientation means he wears many hats at his agile organization. He has become resident IT expert, webmaster, and number cruncher. He pushes his organization to make data-driven decisions, but he lacks the resources, internal buy-in, and spare time to create the processes needed to make this a reality.

After work, Roman is actively involved in his community. He is constantly working on a number of side projects and tends to meet with his collaborators at the Bronx Library. Last year, as he and a few friends were organizing to help low-income families get access to eyeglasses, he asked a librarian to help him find out the number of families living at or below the poverty line. His request was met with shrugged shoulders, and this has always bugged him. He thought libraries were supposed to be sources of knowledge, so why couldn't they help him gather information about his problem?

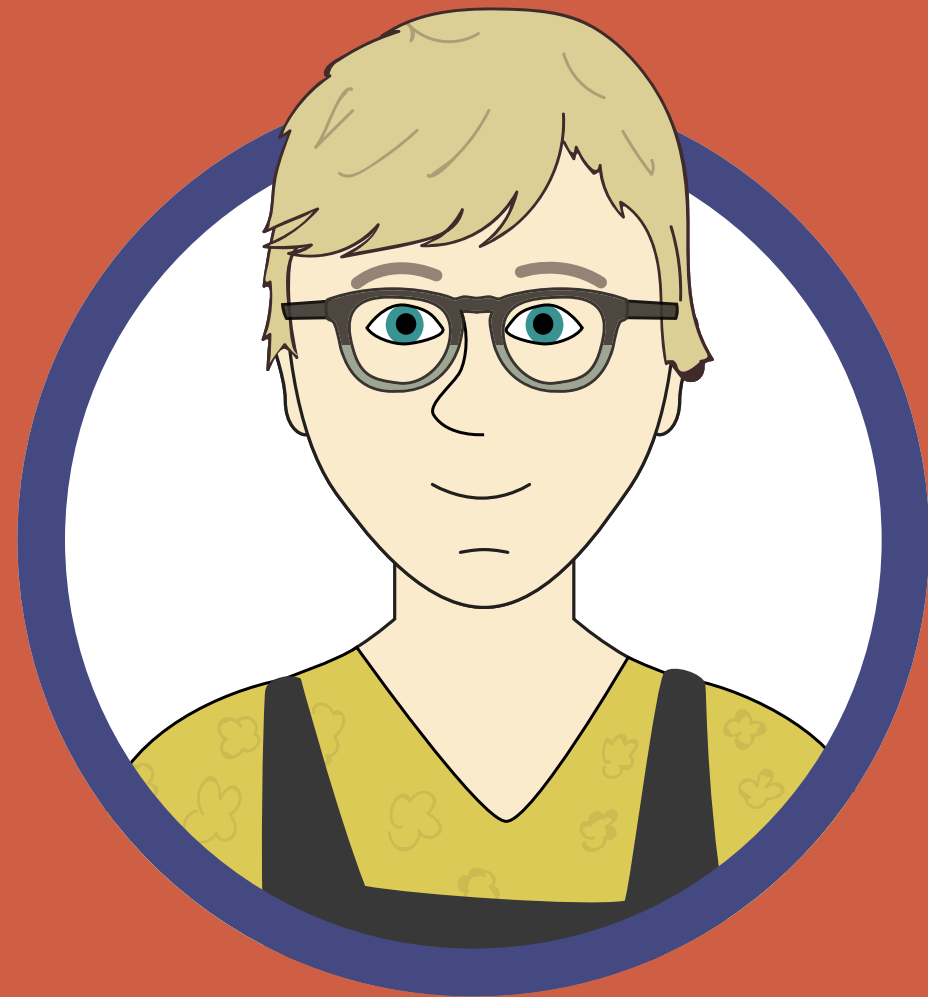
Experience with Open Data

Roman's largest after-work project is an Excel database he built to help both lenders and housing advocates understand the physical condition of buildings all over the city. He has hunted down and cleaned an extensive amount of permit, regulation, and 311 data through a combination of the HPD portal, agency websites, and occasionally the open data portal.

Roman believes in making open source tools that collect and frame raw data in a neutral way. He wants to create tools that bring together essential data sets so that people can use open data to arrive at their own insights and make their own decisions about their communities.

In the past he has brought together community groups and lenders to understand if his database is serving their needs or not—but most of the time he gets feedback much more informally. Often community members will call him up to ask of questions—about Excel, a dataset, or city housing policies.

Although Roman hasn't talked with a manager of a city data set, the operational and bureaucratic barriers that he perceives are associated with city agencies make him skeptical about the reliability of open data. Recognizing that data can be political, he suspects that the even more useful—but politically sensitive—data exists within an agency somewhere, but he doubts he will ever get his hands on it.



JUNE Influential Interpreter

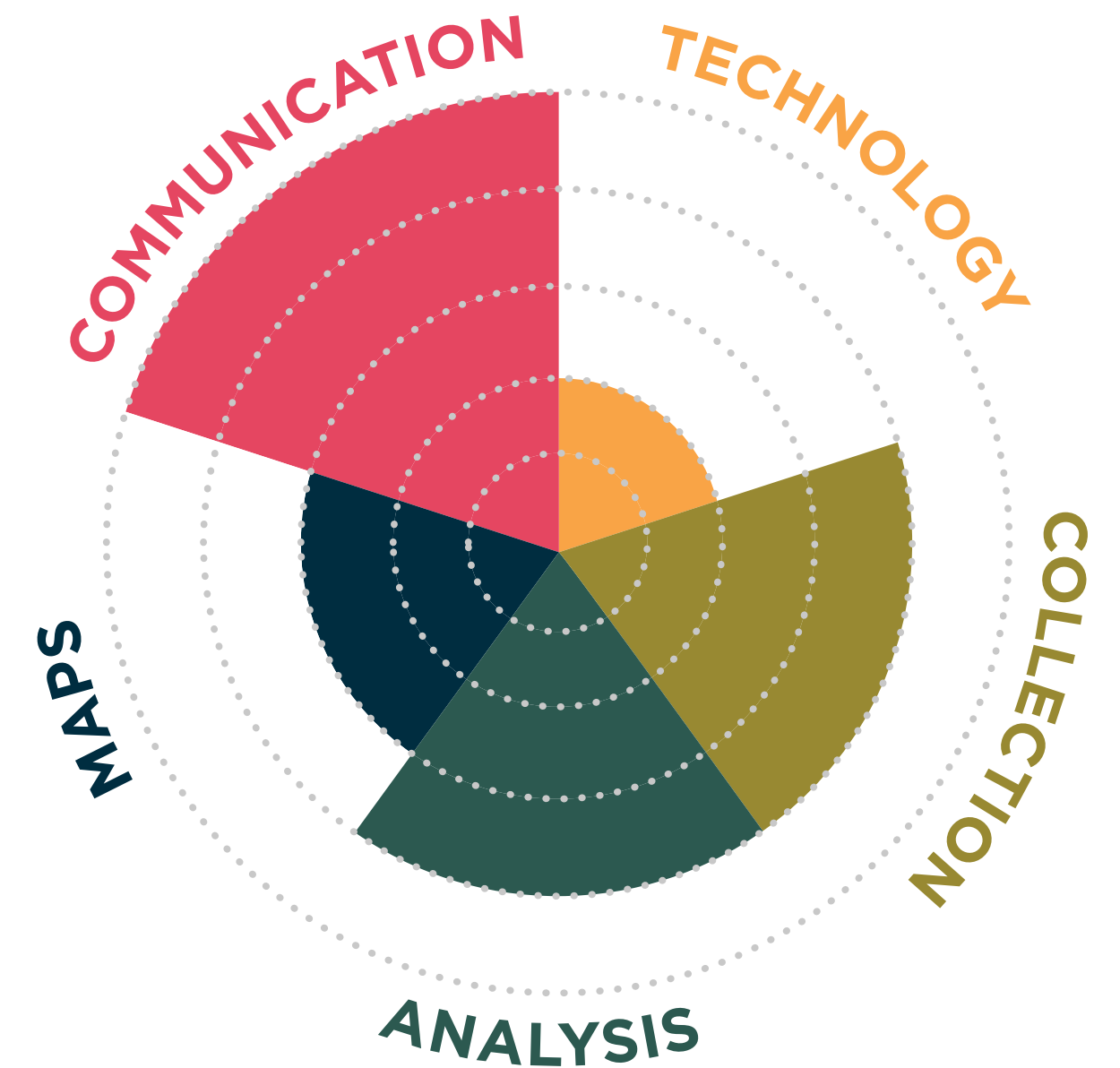
AGE:
28

OCCUPATION:
Policy Associate at an
eviction-prevention
nonprofit

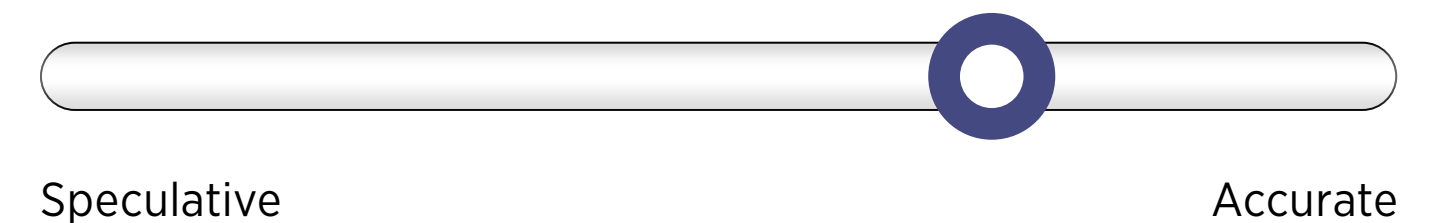
LIKELY LOCATIONS:

Attending a policy speaker series at Civic Hall or New America, at an activist and organizing MeetUp, digging up information for a passion project at a museum archive, attending city council town halls, or working on a data visualization at her local coffee shop.

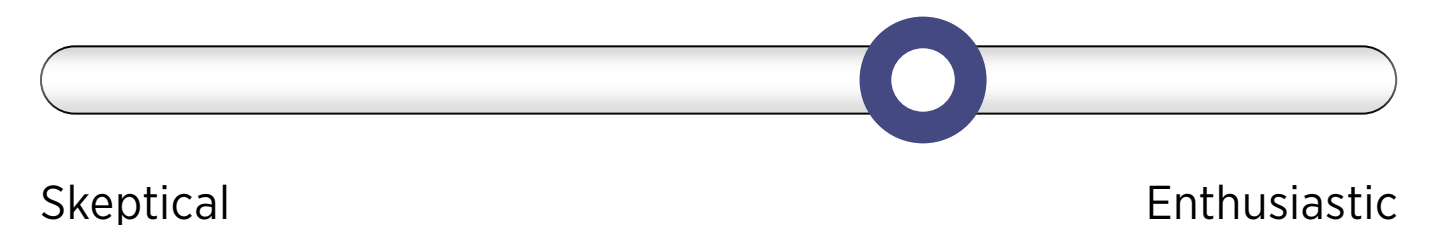
“We don’t want to be the stereotype, and we don’t want to be tone deaf to the communities who have been working on these issues for years. For us, it’s not about how quickly we can get the data out there, but rather it’s about the impact.”



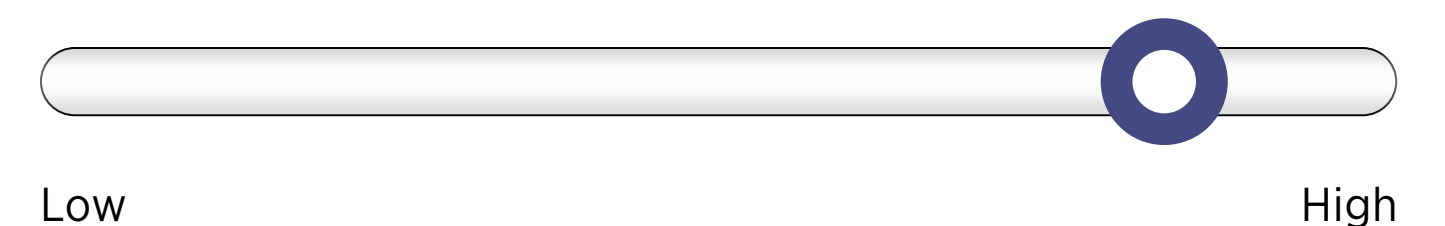
Understanding of Open Data

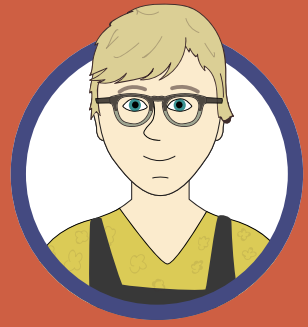


Attitude Towards Open Data



Influence on Community





JUNE

Influential Interpreter

Background:

Shortly after graduating from NYU, June began working at a think tank focused on transportation issues for New York. While she had little experience with data analysis, her new role required her to understand massive amounts of complex data. During her first few months, she relied on coworkers to explain things in non-specialist's terms. Through these relationships she has become fluent in the language of data interpretation.

Excited to apply her skills to an issue closer to her heart, she recently joined an organization supporting single family homeowners facing eviction. Coming from the transportation sector, she was surprised how complicated eviction and demographic data could be compared to the easily measurable and highly precise transportation data.

Half of June's role is focused on translating data sets collected and cleaned by her more technically-focused teammates into meaningful advocacy materials—often in the form of reports, maps, and graphs for her coworkers to take to Albany when speaking with policymakers.

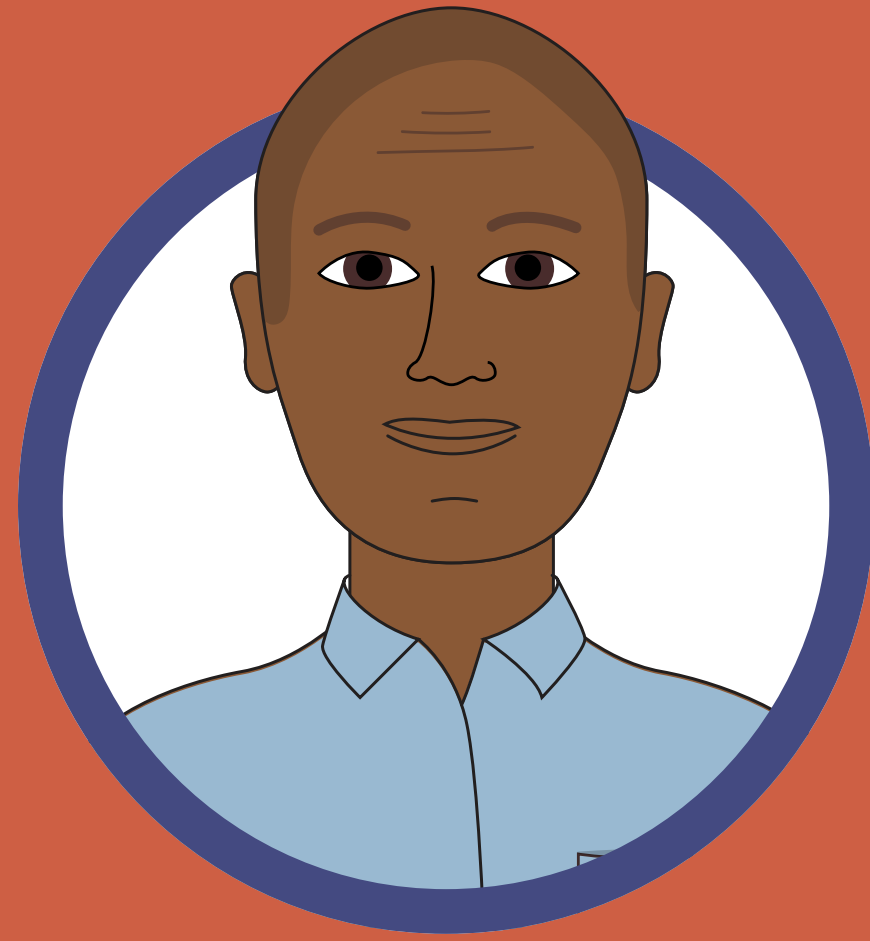
The other half of her role is focused on working with tenants rights organizations and lawyers to ensure that this data is useful to them. These relationships are important to June, since they often don't have the time or skills to dig deeply into public data and they look to her for help identifying trends and recommending policy actions.

Experience with Open Data

June and her organization are motivated by the struggle to safeguard tenants' rights and promote social justice. For them, what matters is not just making quality data available to communities, but ensuring that it is used to fight gentrification. This focus on outcomes has pushed June to think creatively about how to make data both navigable and relevant.

Rather than simply offering partners a collection of maps or graphs, June works with advocates to define how each might use data. Her goal is to build different frameworks that would guide how they make organizing decisions. For example, to help an organization understand the financial means of a neighborhood, she has developed a framework that draws on city-wide salary data and asks users to compare it to a specific set of neighborhood data.

June seems to eat, sleep, and breathe data interpretation. In her spare time, she has found a few friends to help her crowdsource data through Facebook comments and historical archives to map gentrification patterns. She plans to use projections to create an interactive background for a performance art piece.



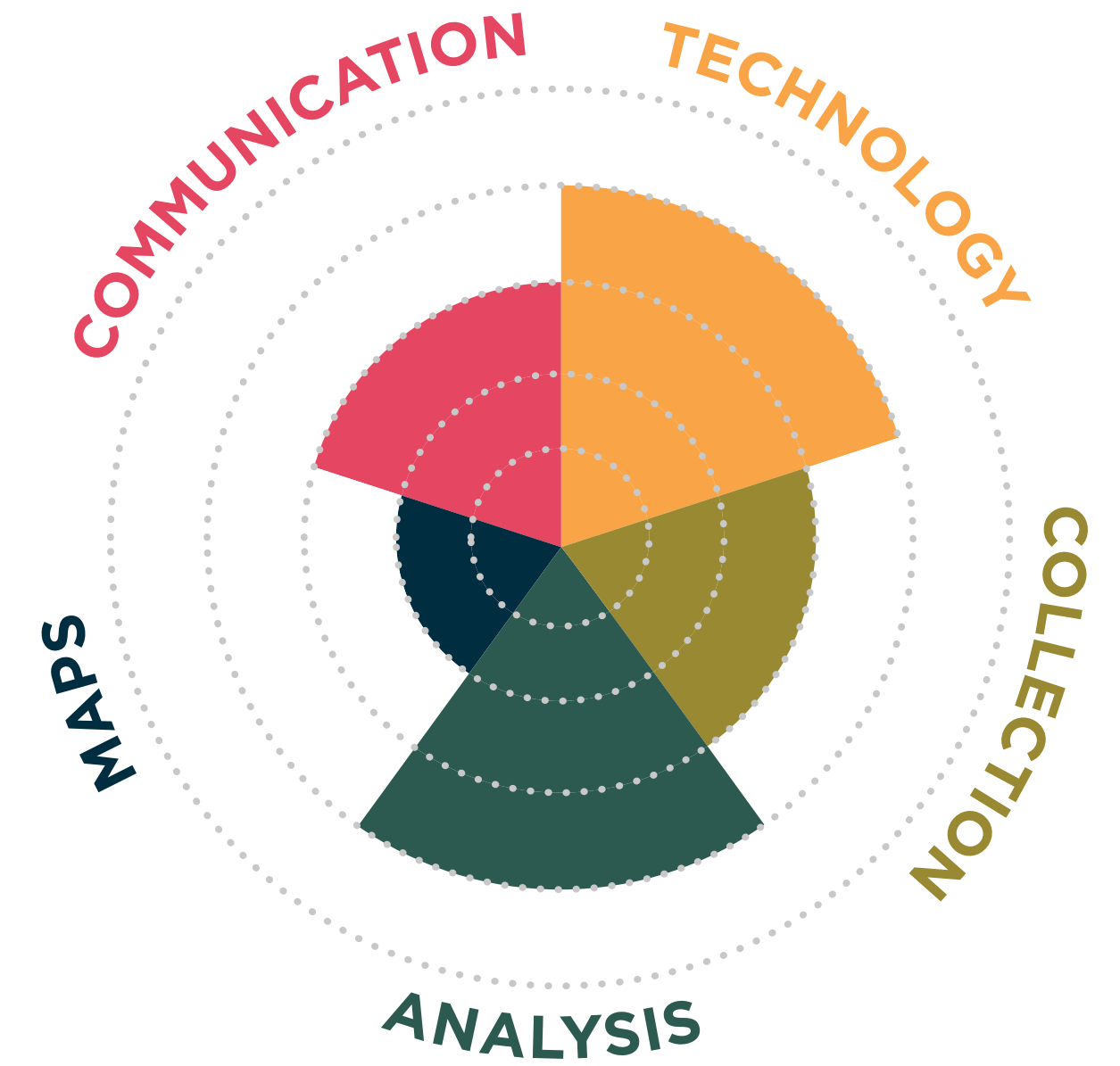
MARCUS Equipped Explorer

AGE:
42

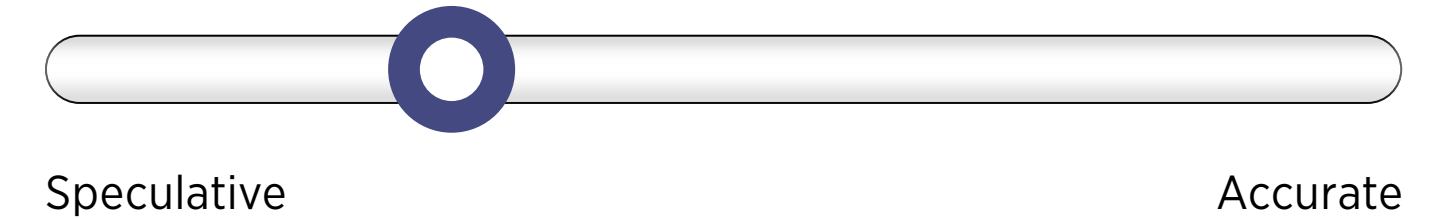
OCCUPATION:
Tech Educator at a college continuing education and workforce program

LIKELY LOCATIONS:
At a parent committee meeting at his son's school, volunteering with a local non-profit organization, participating in a community participatory budgeting program, or tinkering around on a free online education website.

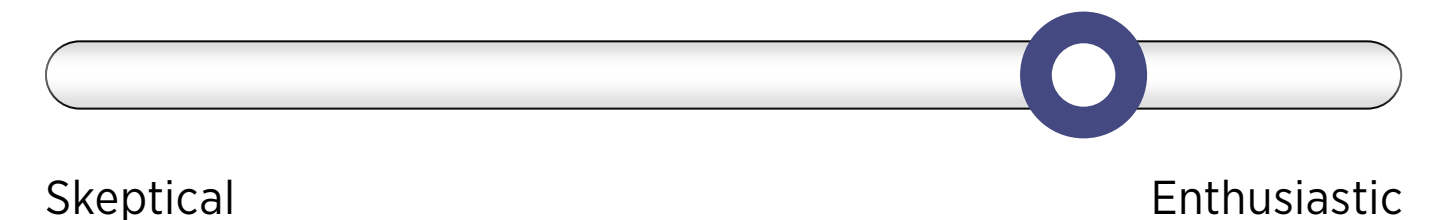
“There is a lot of open data about schools, so we are integrating that into our work. We are still trying to figure out what we want to do with open data, but I think that there is opportunity to improve our work.”



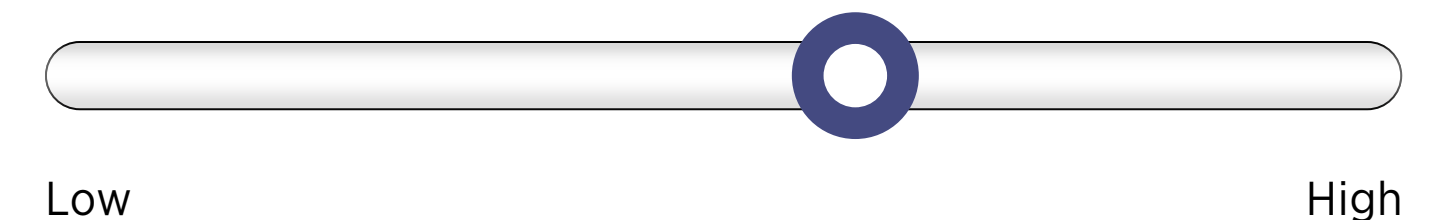
Understanding of Open Data

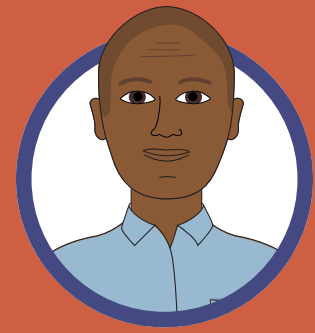


Attitude Towards Open Data



Influence on Community





MARCUS

Equipped Explorer

Background:

From a young age, Marcus' interests spanned biology, physics, and the internet. Moving to New York City shortly after college, he began to teach himself computer science, and started working at an IT company. The company was a great fit for his self-starter attitude and offered countless professional development opportunities, including trips to expensive technology conferences.

After a ten year stretch in IT, Marcus decided to use his love of technology for the social good and began working as a tech educator at a local community college. As he transitioned into the education sector, he missed the formal professional development opportunities from his private sector job. He reached out to his director with these concerns and was able to get sign-off to take a free online class. He recruited a co-worker for motivation and support. Together, during the work day, they set aside time to watch videos and complete assignments.

Outside of work, Marcus is moderately active in his community. He started going to community board meetings when his children were small and he sticks with it when he can. He spends his weekends volunteering with the Gowanus Conservancy—whether working outside or helping with research, he finds ways to make himself useful to the organization.

Experience with Open Data

Marcus has recently become familiar with the open data portal—at first glance he is much more impressed with it than other data sites he has accessed—but hasn't found a way to use it. He browsed through a few data sets and downloaded a subset of 311 data on complaints close to the Gowanus canal—toying with different ideas to support the Conservancy efforts, but nothing came of it.

Although Marcus doesn't have a lot of experience with the open data portal, he has accessed national open data sources in the past. Last month, he attended a Climate Data Jam where he examined climate data, and designed a storytelling project for a non-scientist audience.



SARAH

Busy Bystander

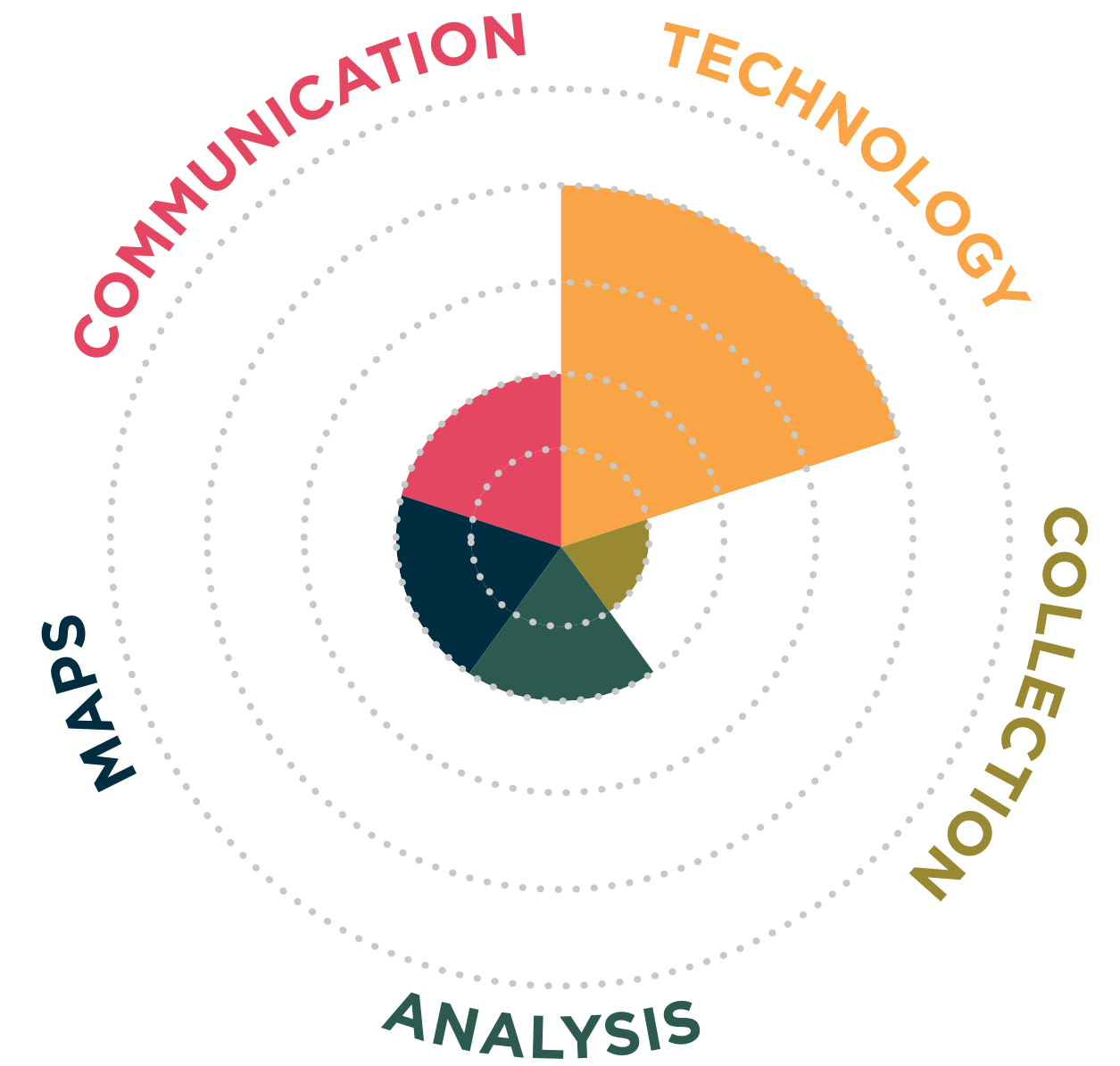
AGE:
22

OCCUPATION:
Student at Queensborough
Community College

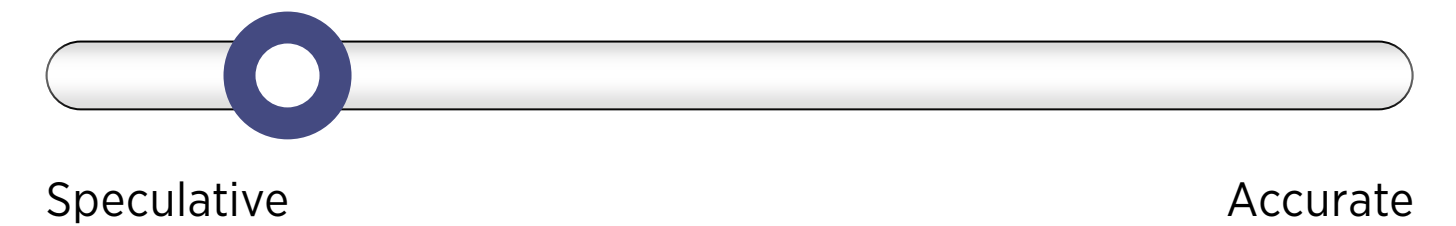
LIKELY LOCATIONS:

At her college campus, on the basketball court in her neighborhood, at home with her family, online on computer science forums, on the bus or subway on her way to and from class.

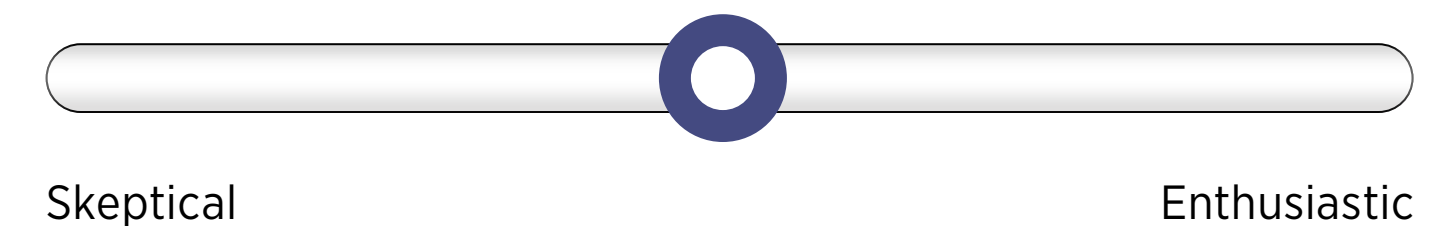
“I like to learn new things, I’m curious. I think open data can make for more informed people.”



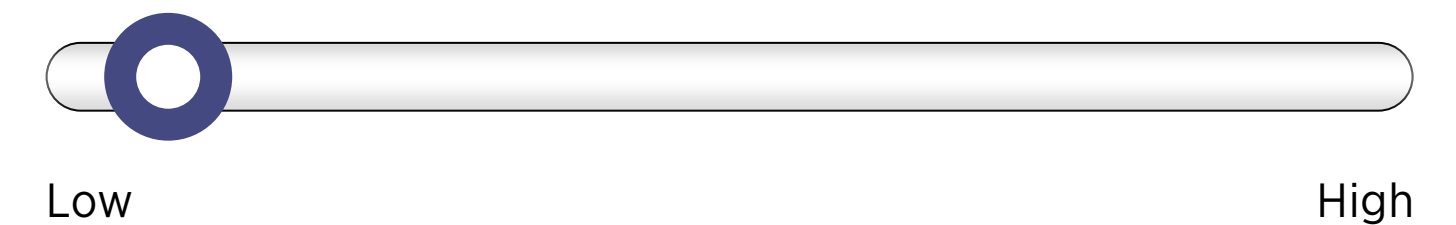
Understanding of Open Data



Attitude Towards Open Data



Influence on Community





SARAH

Busy Bystander

Background:

Sarah is currently majoring in computer science, and dreams of being a developer for Facebook. She grew up in Jamaica, Queens, where her parents encouraged her early interest in physics and math. For as long as she can remember, she has been interested in solving everyday problems through numbers.

Even though she has lived in Queens her whole life, Sarah doesn't feel particularly connected to her community—but she is very close to her family. Her parents run a business selling accessories to local stores and she is always looking for ways to apply what she is learning in school to helping her family. Last year, Sarah and her brother conducted a survey of 60 customers, gathering valuable information about price sensitivity and preferences. Her analysis of the survey results helped her family tailor their business to provide what people wanted.

Experience with Open Data

Sarah mostly accesses data at school and has learned that businesses can benefit from data and research. The value of open data, however, isn't clear to her. She doesn't see a direct application for the data she found on the Open Data Portal, beyond making people more informed.

In one of Sarah's classes, her instructor told her to choose a dataset from the NYC Open Data Portal to use for a data visualization project. She downloaded a dataset on noise complaints in Queens because she remembered some loud construction near her house a few months ago. She thought it was cool that the city had pulled together all that data and enjoyed creating the visualization. But she didn't see what else she could do with it and doesn't imagine herself using the portal except for a school assignment like this one.

In contrast, Sarah recalls another assignment for a physics class that asked her to solve a problem related to her daily life. She took the opportunity to determine the optimal arm angle to throw a basketball. When she was able to get the information to analyze and solve that problem, she was excited to apply it because she could see how it could help improve her game.



ALANA

Community Champion

AGE:

36

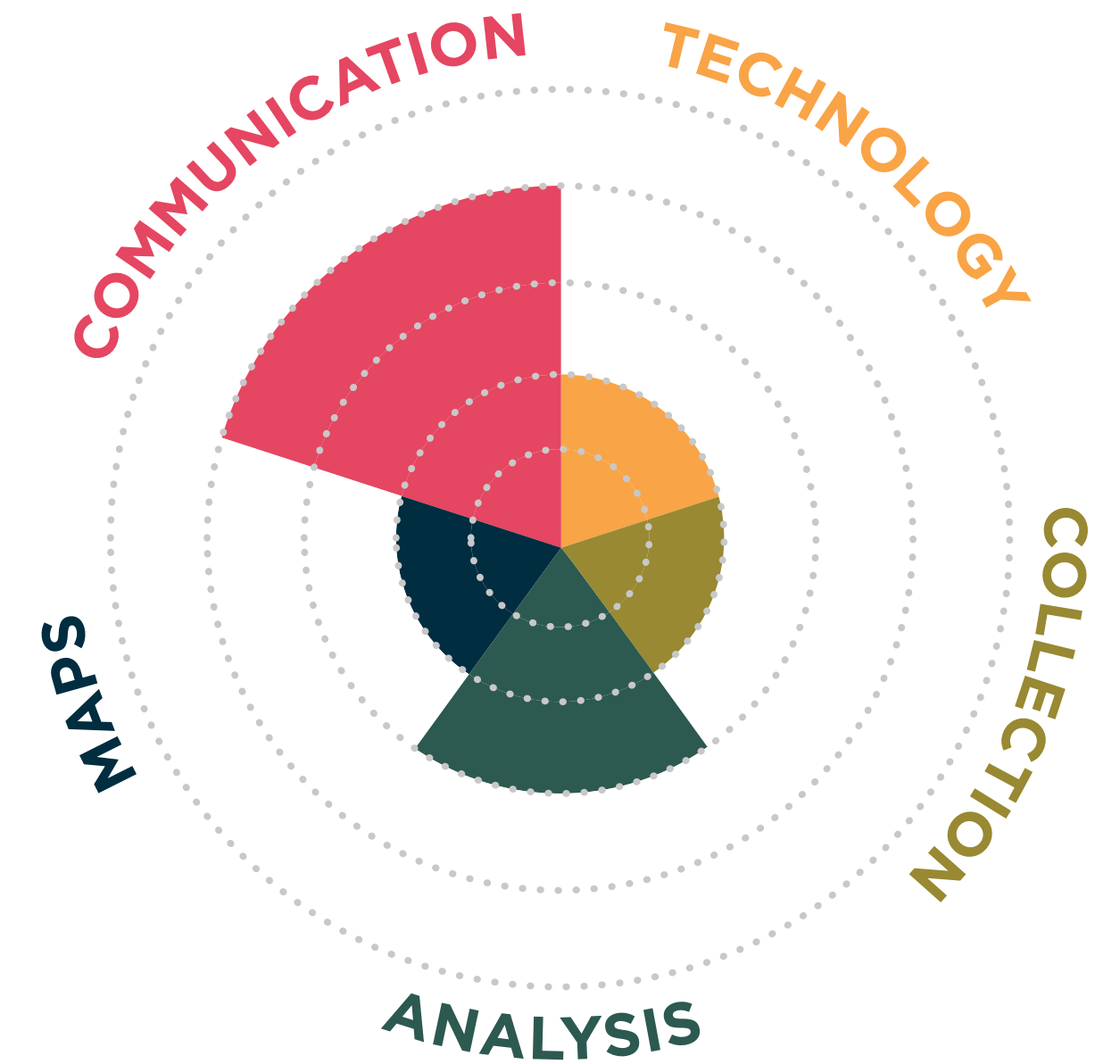
OCCUPATION:

Lead Organizer at an immigrants' rights advocacy organization

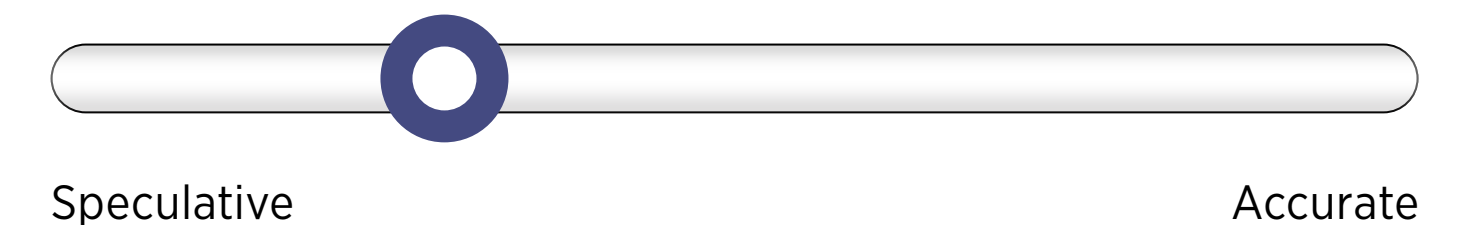
LIKELY LOCATIONS:

At the headquarters of her community organization, the town hall of her local representative, marches and protests outside City Hall, and in public parks, streets and shops of her community.

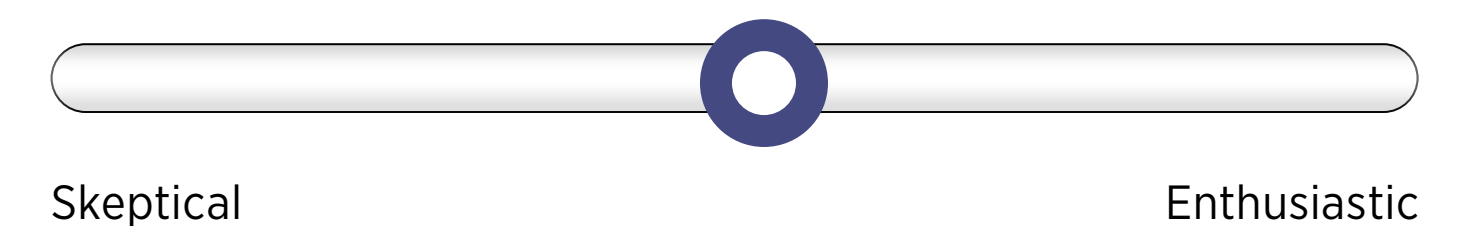
“There’s so much information out there, and if you can connect it to issues you can really help people understand it—but with language and cultural barriers it can be hard to make sure everyone is granted equal access.”



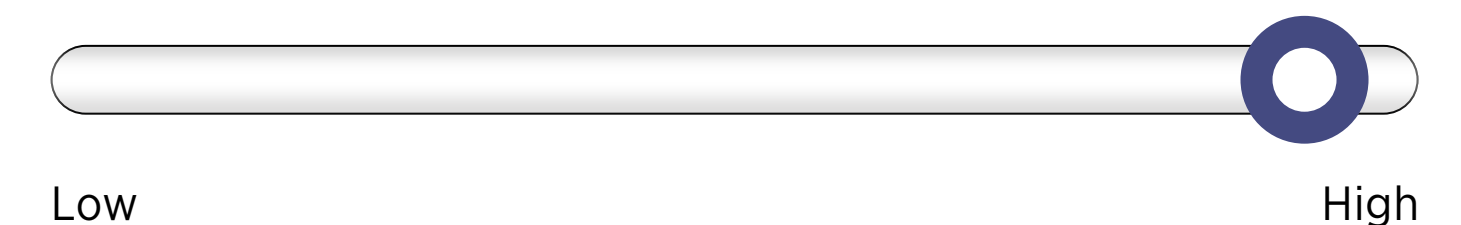
Understanding of Open Data



Attitude Towards Open Data



Influence on Community





ALANA

Community Champion

Background:

The daughter of immigrants from northern Mexico, Alana has always felt strongly about protecting immigrants' rights in the city. Growing up in Borough Park, Brooklyn, she remembers translating street signs and paperwork for her parents.

After graduating from college, Alana worked as a paralegal at an immigration law firm. There, she learned how to use open information databases to collect newspaper articles, court rulings, and other materials that could be used as evidence in court.

Alana is deeply connected to her community. Outside of work, she is dedicated to creating fairer policies for the city's immigrants. Through these activities, she has built a strong network of friends and advocates who collaborate on different community mobilization efforts. She frequently scans the websites of city agencies and city council representatives to learn about town halls she can attend and new legislation that affects her community.

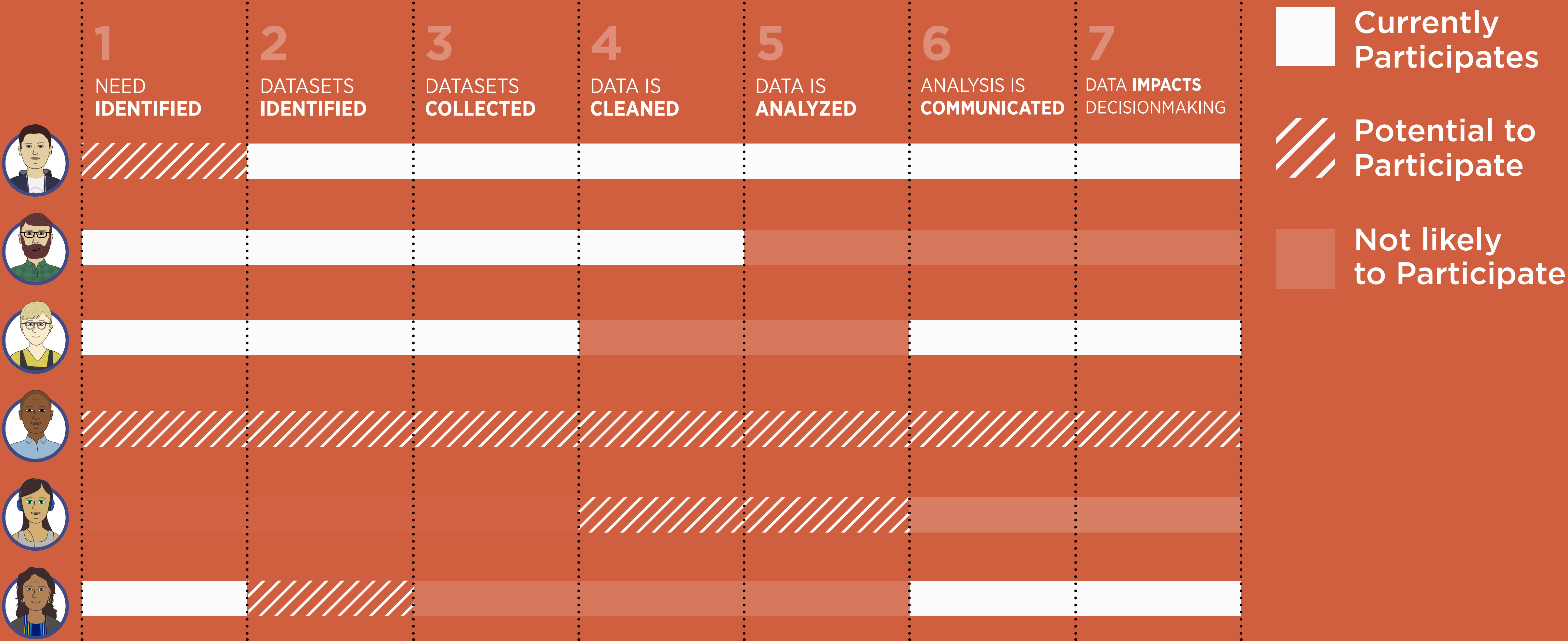
Experience with Open Data

Alana has minimal experience with open data and doesn't find the concept relevant to her work. Though she has experience doing research and finding relevant data sets through work, she is not convinced that so-called open data is a fair or useful resource. Her close relationships with immigrants have allowed her to witness first-hand the many barriers that prevent vulnerable communities from accessing information.

For example, she recently organized a campaign on fair pay that took place in local public schools, because she knows that many families in her community don't own computers or have internet connections at home, so the likeliest place for them to learn about city resources is through their children's schools. That is why Alana and her team focus their energy on producing one-pagers that can be easily distributed in classrooms and taken home, rather than online resources. To pull together and visualize the information she needs, she coordinates with advocates who have design skills. She doesn't have time to learn these skills herself, and draws energy from close collaboration with her network.

Alana is also concerned about data privacy and security issues. Recently, her organization hired a group of volunteers to develop a coding system for their internal database, to protect sensitive information about undocumented immigrants who are members of her organization.

Current Role in the Impact Cycle



Supporting Target Users

Targets:



EXISTING PRIORITY

**Continue improving data quality
to satisfy existing user needs.**

Existing Priority: Continue improving data quality to satisfy existing user needs.



Observations

Experienced users often bypass the portal.

Experienced users utilize existing relationships to data sources.

Experienced users are disappointed with lack of coherent, communicative, and responsive service delivery processes.

Users aren't blind to the politics.

Existing Priority: Continue improving data quality to satisfy existing user needs.



Strategy to Engage

Pull back the curtain on how data is collected, borrowed, and maintained.

Provide integrity and accuracy indicators.

Targets:



PRIORITY 1:

Solidify a strong foundation of supporters who are community-connected.

Priority 1: Solidify a strong foundation of supporters who are community-connected.



Conditions that **Hinder** Open Data Use

Technology paths changing without warning.

Lack of external accountability mechanisms.

Conditions that **Encourage** Open Data Use

Engaging in online learning communities to keep skills up-to-date.

Understanding agency jargon and data collection context.

Building a mutually beneficial relationship with NYC Open Data.

Priority 1: Solidify a strong foundation of supporters who are community-connected.



Strategy to Engage

Position Open Data as a system of resources for a system of people.

SHORT-TERM:

Organize web content according to experience level. Different users are looking for different types of resources.

LONG-TERM:

Since most of these users are going straight to agency websites, work with other agencies to bring their decentralized resources into the Open Data brand.

Priority 1: Solidify a strong foundation of supporters who are community-connected.



Strategy to Engage

Elevate and connect existing work that is creating an impact.

SHORT-TERM:

Hold events that draw community-oriented users to the Open Data brand.

LONG-TERM:

Focus on building relationships that bolster a collaborative ecosystem of open data.

Priority 1: Solidify a strong foundation of supporters who are community-connected.



Conditions that **Hinder** Open Data Use

Protective of community information

Skeptical of open data effectiveness

Disenchanted by current digital information resources

Conditions that **Encourage** Open Data Use

Comfortable with drawing conclusions from research

Advocates for open information and resources

Highly connected with organizers with a wide range of skill sets

Priority 1: Solidify a strong foundation of supporters who are community-connected.



Strategy to Engage

Elevate existing work that is creating an impact to connect users to each other and the NYC Open Data brand.

SHORT-TERM:

Hold events that draw community-oriented users to the Open Data brand.

LONG-TERM:

Focus on building relationships that bolster a collaborative ecosystem of open data.

Target:



PRIORITY 2:

Provide on-ramps for likely users by defining clear use cases and identifying support systems for a path forward.

Priority 2: Provide on-ramps for new users by defining clear use cases and identifying support systems for a path forward.



Conditions that **Hinder** Open Data Use

Looks at the open data portal as an answer box.

Skips over problem-framing.

Minimal experience distinguishing “good” from “bad” data.

Conditions that **Encourage** Open Data Use

Optimistic about the possibilities of open data.

Strong connections to communities ripe to benefit from open data.

Has the necessary technology and data analysis skills.

Priority 2: Provide on-ramps for new users by defining clear use cases and identifying support systems for a path forward.



Strategy to Engage

SHORT-TERM:

Create a welcoming and mentoring brand voice that suggests clear illustrative examples of use.

MID-TERM:

Build “Open Data for problem solving” into trainings and demos.

MID-TERM:

Find ways to connect skilled users with issues by building on existing norms and habits.

Target:



PRIORITY 3:

**Encourage less-likely users
by integrating the “so what”
of open data into existing
educational opportunities.**

Priority 3: Encourage less-likely users by integrating the “so what” of open data into existing educational opportunities.



Conditions that **Hinder** Open Data Use

Sees no value in open data

Misunderstands the meaning of open data

Lacks of civic awareness and/or concern

Conditions that **Encourage** Open Data Use

Looking for the incentives of using open data

Connected to educators looking for opportunities to bring the “real world” into the classroom.

Priority 3: Encourage less-likely users by integrating the “so what” of open data into existing educational opportunities.



Strategy to Engage

SHORT-TERM:

Convene educators to demo open data portal and co-design lesson plans.

MEDIUM-TERM:

Promote the open data brand through public data displays and applications.

LONG-TERM:

Provide lesson modules that use open data as a gateway to other topics for example civic education, social science, statistics, and data analysis.

For more information about NYC's open data initiative contact Adrienne Schmoeker:

aschmoeker@analytics.nyc.gov

For more information about Reboot contact Zack Brisson:

zack@reboot.org

Key for Open Data Skills Maps

This table outlines skill levels observed within the categories of main data skills, surfaced through the research. Use numbers 1–5 within each category to better understand each user persona’s Data Skill Map.

Collection	<ol style="list-style-type: none"> 1. May or may not be able to understand or locate what data sets may be relevant. 2. Understands what kind of quantitative data may be relevant, but may or may not be able to locate it. 3. Has basic knowledge and experience to locate desirable quantitative data sets. 4. Has the experience and/or relationships to seek out hard-to-find data. 5. Has advanced skills necessary to “scrape” data from various sources into a useable format.
Analysis	<ol style="list-style-type: none"> 1. Has a basic understanding of quantitative data. 2. Has experience using quantitative data to understand a problem or issue. 3. Has basic experience using basic programs (i.e Excel) to sort and find themes in normalized quantitative data. 4. Has intermediate experience preparing data for quantitative analysis; including cleaning, normalizing, and standardizing. 5. Has extensive experience using advanced applications to manipulate data.
Mapping	<ol style="list-style-type: none"> 1. Has a basic understanding of maps. 2. Has basic experience using maps to surface and understand issues. 3. Has intermediate experience using maps to surface and understand issues. 4. Has experience mapping relevant data sets to geographic areas using common computer applications (i.e. Carto). 5. Has skills and significant experience building maps.
Technology	<ol style="list-style-type: none"> 1. Has experience using basic technology including internet browsing and searching. 2. Understands basic technological structure. May not have significant experience, but can translate ideas to technology stakeholders. 3. Understands how to utilize components of technology. Possesses basic coding skills. 4. Has experience in computer programming. Understands how to contribute and collaborate with version control software like Git. 5. Is fluent in computer programming skills. Capable of building complex applications using a variety of computer languages.
Communication	<ol style="list-style-type: none"> 1. Has a basic understanding of how to explain information to someone without little topical experience. 2. Has experience organizing and presenting information to people with little topical experience. 3. Has basic experience organizing and presenting quantitative information to people with little quantitative experience. 4. Has extensive experience creating communicating insights from quantitative information, including creating visualizations through analysis software (i.e. Tableau) written reports that communicates insights from data analysis to the appropriate audience. 5. Regularly trains infrequent or unfamiliar quantitative information consumers in how to use and understand data analysis.