

Giving City Residents Greater Access to Local Quality-of-Life Data

Andrea Sepúlveda-Vargas^{1,2}, Vishnu Nair², Brian Smith²
University of Puerto Rico - Mayaguez¹, Columbia University²

Abstract: Cities produce large amounts of data such as public service request data and local government meeting records. However, much of this data requires technical knowledge or a time investment from the user to access, making this data accessible only to a limited group of people. We believe there is an opportunity to make this information more accessible to everyday community residents. In this project, we explored how we could process publicly available data such as that from NYC Open Data and NYC Community Boards, to make it more transparent to residents. Our findings from Harlem, New York reveal patterns within the data that point to possible hyper-local issues. Future work points to user-facing systems that compile and expose these patterns in an accessible way, leading to greater community awareness and more consistent action from governments. To this end, we present high-fidelity prototypes for a mobile app that visualizes these community-specific concerns.

Methods: We conducted a series of group brainstorming sessions to investigate ways we could enhance the accessibility of datasets within NYC Open Data and issues discussed during Harlem Community Board meetings. As part of this, we reviewed existing approaches that give city residents access to information about their communities and local government efforts.

We used 311 complaint data and community board meeting minutes to understand what is happening within the community — focusing specifically on Community Boards 9, 10, and 11, which cover Harlem. We extracted 311 complaint data from the corresponding dataset on NYC Open Data. To analyze community board meetings, we fed meeting minutes from 2024 through GPT-4o, which summarized each meeting, extracted discussed issues, and associated each complaint with a physical location in Harlem. We then used Kepler.gl to plot 311 complaints and community board issue locations on an interactive map of Upper Manhattan, which we then explored as part of our analysis.

Results: Our findings revealed that effectively utilizing NYC Open Data requires significant technical skills and data science expertise. Existing tools are highly specialized, addressing only a limited number of specific datasets and use cases. Also, we found that information about Community Board discussions is scattered across multiple websites and formats, making it difficult to access and understand. This fragmentation decreases transparency for everyday residents.

Our study uncovered hidden patterns within the publicly accessible data, particularly focusing on areas with

clusters of 311 complaints. As an example, *Figure 1* visualizes the densities of 311 complaints made in 2024. We found the highest number of complaints (3,314) were filed at the location of bright red bar toward the left — surprisingly, that bar is right next to Columbia University (115th and Broadway)! The second-highest number of complaints, represented by the orange bar with a total of 2,276 requests, is at 140th Street and Adam Clayton Powell Jr. Blvd.

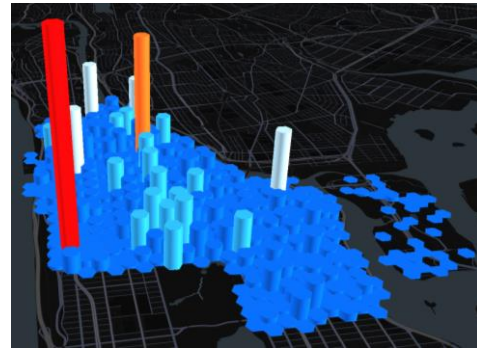


Figure 1. 311 complaint densities in Harlem during 2024.

These patterns can help both residents and governments quickly identify community issues and understand the urgency of these. In the case of the area around Columbia, the most frequent complaints centered around illegal parking (1,867) and mobile food vendors (923). Similar trends could be found within other areas of Harlem.

Prototypes: Our prototype features an intuitive interface that improves user accessibility to information relevant to their environment. We generated a platform that exposes data visually to make it more understandable and useful to the communities. We used the Upper Manhattan Community Board map for better user visualization. Each dot presented in the map, *Figure 2*, represents a request made by a resident in that location. The color of the dots represents the agency name in charge of that specific complaint.



Figure 2. First Harlem Community application mockup (left) versus last version (right).

Once the user selects which Community Board interests, they will be able to dive deeper into the information of each

request. The information provided includes the agency name, complaint type, description, location type, street name, community board number, and the created date. Also, the user will be able to manage the complaint request date time with the slide below the page, as shown in *Figures 3-4*.

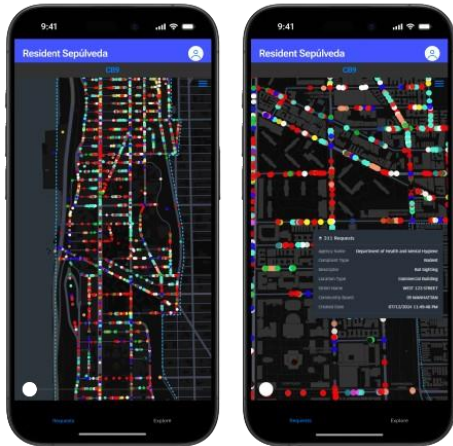


Figure 3. Visual representation of 311 complaints in West Harlem (left) with the complaint information box (right).

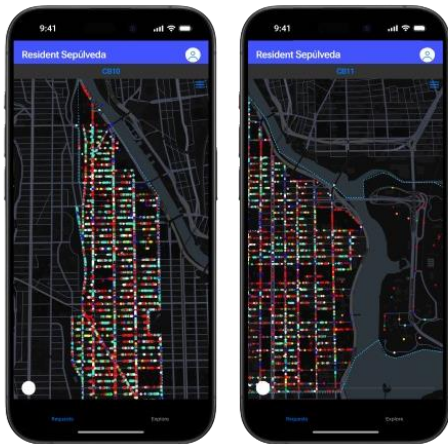


Figure 4. Visual representation of the 311 complaint requests in Center (left) and East Harlem (right).

The “Explore” tab will feature two sections: Community Boards and Petitions, as shown in *Figure 5*. The Community Boards page will display an Upper Manhattan map with dots representing each issue category discussed in Community Board meetings, accompanied by relevant information. The Petitions page will have a submission option where residents can submit complaints. Other residents who have the same concern can vote to support it.

Increasing the votes for a petition will lead to LLM agents tracking relevant information that supports each complaint with real data. As an example, *Figure 6* shows the Supporting Evidence section which has the source and the link to the raw data found. This information will help government leaders better understand the voice of the community and take action with real evidence.

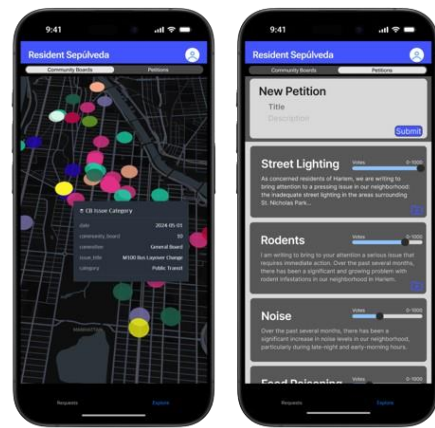


Figure 5. Issue categories discussed in the Community Boards 9-11 during 2024 (left). Petition submission section for Harlem residents and examples of submitted petitions (right).

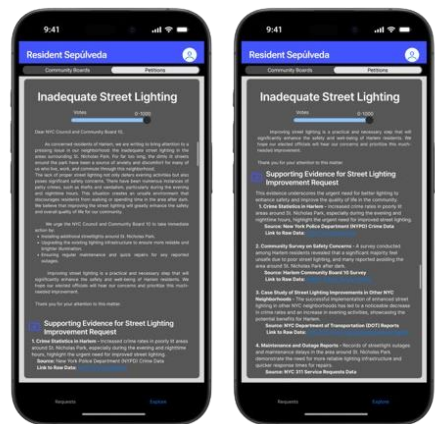


Figure 6. Petition section for Harlem residents with LLM supporting data section.

Conclusion and Future Work: This study contributes to a broader understanding of how we can better use public data to increase awareness within and empower communities. Current means of accessing open data require time and technical acumen. To address this, future work should focus on developing a more integrated and user-friendly platform that compiles public information within a single system and supports concerns raised by the community with relevant data. Overall, this study promotes the importance of truly accessible open data to address significant resident concerns — all so that the community can have a better quality-of-life.

Acknowledgments: I would like to express my sincere gratitude to the Computer Enabled Abilities Lab (CEAL), for their guidance and support throughout this research. This study was supported by the Amazon Summer Undergraduate Research (SURE) program.