A Project Report

Leveraging Citizen Science for NYC’s Real-Time Monitoring and Communication Using Noise App as Case Study

For

TOWN+GOWN: NYC, NYC Mayor's Office of Climate & Environmental Justice, and NYC Department of Environmental Protection

From

New York University, Tandon School of Engineering
Department of Technology, Management and Innovation

Capstone Project Team
Charlie Chung
Pragathi Thoranagadde Ravishankar
Zihao Gao

December 12, 2025

TABLE OF CONTENTS

1. Abstract

Urban noise is a persistent environmental challenge in New York City. While tens of thousands of complaints flow through the 311 app each year, the existing monitoring systems still leave gaps in real-time coverage and public participation. Building on the prior Town+Gown project that designed a citizen science smartphone app for reporting urban noise, this capstone project investigates how NYC can validate and operationalize citizen-generated environmental data using the noise app as a case study.

Through preliminary research in the form of literature reviews and primary research in the form of user interviews with various experts, the project examines key issues such as data quality, sampling bias, privacy, community engagement, and agency workflows when integrating citizen science into official monitoring systems. A proposal for the validation direction is then presented to assess the reliability and usability of data, focusing on criteria such as accuracy, contextual metadata, and aggregation strategies. The project also dives into design considerations and constraints, eventually with suggestions for a future unified, public-facing dashboard that could integrate other environmental hazards beyond noise.

These suggestions are based on the literature review, expert interviews conducted, and through exploring existing city pilots such as FloodNet. These findings, when taken together, offer Town+Gown: NYC, NYC DEP (New York City Department of Environmental Protection), and NYC MOCEJ (New York City Mayor's Office of Climate & Environmental Justice) a pathway to turn citizen-sourced data into trusted, actionable information while strengthening community engagement and transparency in environmental governance.

2. Acknowledgements

We would like to express our sincere gratitude to Town+Gown: NYC, DEP, and MOCEJ for sponsoring this project and providing continuous guidance throughout the semester. We are especially grateful to Terri Matthews, Mark Page, Jr., and Hayley Elszasz for their time, insights, and support in shaping our research direction.

We would also like to thank the experts from FloodNet, DOHMH, NYU Tandon, CUNY, and other experts who generously shared their time and insights during our interviews. Their contributions were essential in helping us understand the opportunities and challenges of citizen science–based environmental monitoring in New York City.

Finally, we would also like to express our appreciation to Professor Christopher Policastro and our teaching assistant Errol Elbasan for their guidance, feedback, and support throughout the development of this project. We would like to thank all team members for their collaboration and dedication throughout the semester.

3. Introduction

Noise pollution has become a familiar part of daily life in large, busy cities. As urban spaces grow denser and activity continues around the clock, sounds from traffic, construction, and commercial areas blend into a constant background that can affect how people live, work, and rest. Over time, unmanaged noise can contribute to stress, fatigue, and other challenges that impact overall well-being.

In New York City, noise is one of the most common topics reported through the 311 system. Residents frequently raise concerns, but the city's current monitoring tools, especially fixed sensors, do not cover enough neighborhoods to show the full picture of where and when noise problems occur. Because sensors are costly and limited in number, the city cannot know enough, and community members can only use 311 to report. To begin addressing this gap, a previous MOT Capstone project worked with NYC DEP to create a prototype mobile app that allows residents to record noise events directly from their smartphones.

Building on the earlier work, this project focuses on how the city might use citizen science data in a more structured and reliable way. By speaking with experts in the field, we explored what it would take to validate this type of data, gather community engagement, and how it might eventually feed into a broader platform in the future for environmental information.

4. Project scope

This capstone project was undertaken in collaboration with Town+Gown: NYC, DEP, and MOCEJ. The client team seeks to map out the road toward validating and incorporating citizens-generated noise data into New York City's environmental monitoring and decision making. Whereas the previous MOT Capstone developed an operational smartphone app prototype for noise data collection, this work focuses on research, validation, and the strategic application of citizen-generated data within existing public-sector workflows..

4.1 Research objectives

- Assess the method of data validation: Test technical and process means for validating the correctness, trustworthiness and interoperability of citizen-sourced noise data.
- Point out data integration issues such as data bias, data privacy, metadata, and agency workflow integration.
- Expert engagement: Conduct interviews with academic, government, and private-sector experts to understand the barriers and potential of citizen science in official environmental monitoring.
- Recommend implementation: Provide the client team with concrete guidelines on how to implement the citizen-generated noise data in a secure and scalable way.

4.2 Key deliverables

- Interview synthesis: Synthesis of insights from interviews with subject matter experts across academic, government, and private sectors.
- Data validation framework: A synthesized set of principles, criteria, and mechanisms for assessing the quality and practical usability of citizen-generated noise data, informed by prior research and expert interviews.
- Final report and recommendations: The final report provides a comprehensive account of the research findings and outlines recommended next steps for client teams to support the future development of the NYC Noise Monitor App, launched by NYC DEP in November 2025, as discussed in Section 8 below.

5. Literature review

5.1 A Smartphone-Based Crowd-Sourced Database for Environmental Noise Assessment *(Picaut et al., 2021)*

Picaut et al. introduce NoiseCapture, which is a crowdsourced, smartphone-based platform for environmental noise assessment that was created as part of the Noise-Planet project with a strong open scientific mindset that includes open source code, open data, and open access publications. The constraints of conventional strategic noise maps and stationary observatories, such as being sparse, expensive, frequently falling short of capturing subtle temporal, and geographical fluctuations in urban noise, are outlined by the authors. When, however, paired with suitable metadata and quality control procedures, citizens' smartphones can function as dense, mobile sensing devices, producing vast amounts of in situ measurements.

Over the course of three years, NoiseCapture was downloaded over 160,000 times and generated over 260,000 measurement "tracks" and around 60 million one-second acoustic data points globally, illustrating the scope of environmental noise mapping that citizen science can do. The fundamental design of NoiseCapture addresses issues that are pertinent to the context of the Fall 2025 Noise App design case study. While walking or standing, users start measuring "tracks," during which the software locally calculates third-octave spectra and A-weighted sound pressure levels on the phone without storing raw audio. GPS location, time, user speed, and device orientation are added to each one-second measurement point. Following the track, users can choose to tag contextual data like environment type (such as "road," "rail," or "indoor"), subjective pleasantness, or flags (like "test," "rain," or "wind.") A public map that aggregates observations into hexagonal cells and displays metrics like LAeq and dispersion statistics over space and time is created by uploading all contributions to a central PostGIS database. It follows that while basic decibel reporting is a reasonable place to start for the new NYC Noise Monitor App, deeper metadata will be essential for future validation and policy use. For example, the app should, at the very least, record the device model, GPS accuracy, motion state (walking, stationary, in-vehicle), and whether the user marks the report as a "test" versus a genuine complaint scenario.

The authors of this paper also discuss data quality and constraints, which relate to developing a validation framework as opposed to being merely a data gathering tool. Picaut et al. list a number of issues with the raw database: a significant portion of contributions are very brief "test" tracks that were recorded by one-time users; many tracks have missing or inaccurate GPS; some measurements come from devices with unknown or subpar microphone characteristics; timestamps and speeds can be inconsistent. Simultaneously, the authors demonstrate that the majority of the data is contributed by a very limited number of smartphone models and manufacturers, and that active contributors exhibit different behavior from casual users. They conclude that systematic filtering by device type, app version, calibration state, GPS quality, and contributor behavior is necessary for useful database use. This implies that NYC DEP and MOCEJ should move from the question about "are citizen data valid" to the question, "under what conditions, and with what filters, are these data reliable enough to inform investigation, mapping, or policy discussions."

The study finally concludes by highlighting the significance of contributor behavior and community dynamics, pointing out that a tiny percentage of active users provide a disproportionate amount of high-value data. The authors recommend that citizen scientific noise platforms identify and assist these "power users," for instance by providing customized advice, comments, or gamified rewards, while simultaneously creating user interfaces that reduce mistakes made by infrequent users. This suggests that all reports should not be handled uniformly by the app that is designed and any associated 311 protocol in NYC. Alternatively, concepts like contributor history, consistency over several reports, agreement amongst independent users in the same area, and match or mismatch with fixed sensors or modelled noise maps can be incorporated into the validation framework. When combined, these observations from the article can become a useful guide for NYC, emphasizing that citizen-generated noise data can be highly valuable if it is collected, labeled, and filtered through a deliberate, well-defined validation framework.

5.2 iCal - Intervention-free Calibration for Measuring Noise with Smartphones *(Zhu et al., 2015)*

In the framework of citizen science for environmental noise monitoring, data quality and sampling bias are among the most important issues. This research offers essential empirical support and a technical approach to reveal and treat these problems:

- It validates the relevance of data quality as a necessary premise for the viability of citizen science, and therefore also demands a straightforward confrontation of the accuracy gap between mobile and professional devices.
- It offers a practical data validation framework by using algorithmic models to perform automated, device-specific calibration, which is the key to transforming citizen data into trusted information.
- It informs strategies to manage sampling bias, suggesting that crowdsourcing can correct for hardware-induced bias, while the mobility of citizens is an inherent advantage in overcoming geographical sampling bias.

This research first conducts controlled experiments finding that the principal difference in data quality between consumer smartphones and professional devices lies in a lack of calibration, illustrating that measurement errors may be as much as 15 dBA across varying brands and models with the uncalibrated smartphone, and there are significant and nonuniform variations across phones. It suggests that amateur, unvalidated sensor phone data measurements cannot be scientifically ‘swapped’ with official readings and do not have sufficient precision to be used for environmental monitoring and policy actions.

Nonetheless, there may be ways to deal with this quality deficit by providing systematic calibration. The research specifically found that magnitude error is large enough for the relationship between the measurement provided by a smartphone and the actual noise measured by professional meters to be stable over time (i.e., to follow a linear model), and it is simple to establish the relationship by establishing a linear calibration model (i.e., estimate the relationship from a subset of ground truth and propagate such relationship to the entire dataset.). Once a device-specific offset can be computed for every phone (after the process of offsetting from a device to another one), the calibration procedure can account for measurement errors in a systematic way. The calibration process for node-based approach is inspired by a robust estimate of a common quiet noise background (approximately 36.5 dBA) present in most indoors environments, as a “reference weight” that is used to derive a device specific calibration parameter per individual device. The calibration procedure was found to reduce the mean error from more than 10 dBA to around 3 dBA on average, significantly enhancing the trustworthiness of data collected by citizens.

On sampling bias, this research leads to the following implications. Professional sensors are static and expensive and therefore often have sampling biases due to limited geographic coverage. Citizen science is mobile, which is in principle able to compensate for coverage gaps. Crowd-sourcing-based calibration module leverages the mobility to that advantage. It turns a loose network of calibrated devices into portable ”calibration sources”, and sharing model-specific calibration parameters mitigates the systematic bias in the measurements caused by heterogeneous hardware. Hence, it is possible to enable data from different users using different phones, from different places, to be interpreted consistently, standardized, and therefore more comparable.

5.3 SONYC: A System for the Monitoring, Analysis and Mitigation of Urban Noise Pollution *(Bello et al, 2018)*

Urban noise governance relies heavily on complaint-driven systems, such as New York City’s 311 platform, yet these systems are known to suffer from spatial, temporal, and demographic biases that limit their effectiveness in reflecting actual noise conditions. To address these gaps, Bello et al. (2018) introduced the SONYC (Sounds of New York City) initiative, a cyber-physical system combining low-cost acoustic sensors, machine listening, and spatiotemporal analytics to generate continuous and source-specific noise data across the city.

This study demonstrates substantial discrepancies between citizen complaints and objectively measured noise levels. For instance, in the context of after-hours construction, sensors detected noise events in 94% of complaint cases, while only 2% resulted in official violations due to delays and limitations in the inspection process (Bello et al., 2018). These findings highlight the limitations of complaint-based monitoring and the need for alternative data collection approaches capable of providing real-time, high-resolution observations.

The SONYC framework offers several insights relevant to emerging citizen-science approaches in environmental monitoring. First, its validation pipeline (combining calibrated sensors, audio classification algorithms, and cross-checking across spatial and temporal datasets) illustrates how cities like New York can ensure data credibility before bringing information into public-facing platforms. Second, its visualization system, Urbane, demonstrates the feasibility of combining heterogeneous environmental data streams into a unified dashboard that can support public transparency and agency-level operational decision-making. Third, the enforcement gap identified by SONYC indicates that technological innovation should be paired with policy and workflow adaptation to translate environmental data into actionable regulatory outcomes.

Together, these contributions set SONYC as a foundational model for understanding how local governments can leverage distributed sensing to supplement traditional monitoring infrastructures. SONYC also helps establish a broader scope for the capstone project, which investigates how the government could incorporate citizen-generated noise data, collected via a mobile phone application, into a validated, citywide dashboard. By illustrating both the promise and limitations of sensor-based monitoring, SONYC provides essential context for exploring the design of data validation procedures, dashboard architectures, and governance frameworks that can support citizen science as a reliable component of noise management systems.

6. Methodology

Understanding how citizen-generated noise data could be verified, contextualized, and ultimately included into NYC citizen science-based procedures was the foundation of the capstone's methodology. To achieve this, the team adopted a mixed-methods approach combining literature-based analysis, stakeholder interviews, and multi-source benchmarking. A thorough analysis of academic and technical literature on citizen science noise monitoring, including smartphone-based sensor systems, calibration techniques, privacy concerns, sample bias, and earlier municipal pilots, was the first step in the research. This helped establish a technical foundation and highlighted the core challenges of data accuracy, device variability, and contextual metadata, that the validation framework would need to address.

The team conducted semi-structured interviews with various experts in academia and in practice with experience in crowdsourcing, environmental sensing, urban informatics, and municipal governance, building on the ideas found in the literature. Professors with expertise in crowdsourcing sensing, urban mobility, and noise monitoring were among the interviewees, along with representatives from related city initiatives like FloodNet. These discussions revealed design issues for any future environmental data dashboard, as well as practical limits including data filtering requirements, sampling limitations, and inter-agency communication issues.

The team used recurrent themes found throughout conversations to synthesize interview notes rather than using a formal coding structure. Data accuracy, calibration viability, user behavior, privacy, enforcement limitations, and inter-agency coordination were among these themes. The team was able to determine where technical and practical considerations aligned, such as the significance of metadata, GPS accuracy, and device awareness, and where additional investigation might be required by comparing the recurrent patterns in stakeholder feedback with the difficulties mentioned in the literature.

These combined methods of literature review, experts interviews, and high-level examination of related initiatives, enabled the team to identify key considerations for NYC's noise data validation needs. In the context of the new NYC Noise Monitor app, the ensuing insights directed the creation of preliminary suggestions detailing minimal needs and future prospects for enhancing trust, accessibility, and actionability of citizen-generated noise data.

7. Takeaways from interviews

7.1 Insights from academic experts

In the opinion of both professors, Dr. Charlie Mydlarz and Dr. Hamidreza Norouzi, citizen science data has the potential to be a very useful complement and augmentation to official urban monitoring systems. Dr. Mydlarz is a Research Associate Professor at the Center for Urban Science and Progress (CUSP) at the NYU Tandon School of Engineering and the Music & Audio

Research Laboratory at NYU Steinhardt. He has many years of experience in urban noise pollution control and the impact on the sound environment. Dr. Norouzi is a Professor of Civil Engineering at the City University of New York (City Tech, CUNY) and the CEO and Co-Founder of Eco Rising Solutions, a geospatial and AI-driven environmental technology company. He has conducted in-depth research on environmental pollution. One key lesson from them is that the primary challenges in making citizen-generated data useful include clearly framing stakeholder expectations, recognizing the limitations of the data, and designing resilient systems that achieve acceptable data quality while maintaining participant incentives and data privacy.

Dr. Mydlarz highlighted both the value and the limitations of citizen-science data. On the value side, he noted that citizen-generated data provides broader spatial coverage than fixed sensors alone. Because people move around with their phones, they can capture data in locations that are difficult or expensive to monitor permanently, effectively extending the city's sensing network.

However, several limitations remain. He emphasized variable data quality, poor temporal resolution, and challenges in recruiting and motivating participants. For data quality, it is not feasible for the public to professionally calibrate their phone microphones, which makes the data unsuitable for reporting precise absolute noise levels. Instead, it is better suited for identifying general ranges or categories, such as a "loud" band. Regarding temporal resolution, citizen-science data cannot match the continuous data stream provided by fixed sensors. Finally, because project success often depends on maintaining a large and active user base, sustaining reliable user participation and behavior is difficult.

Based on these advantages and limitations, Dr. Mydlarz believes that citizen-science data is best applied in early screening and trend tracking, where it can provide useful assistance in decision-making rather than serve as a precise measurement tool. Capturing short audio clips is especially valuable because it allows inspectors to identify the noise source (e.g., pile driving vs. HVAC), which is often as important as the noise level itself. He also noted that privacy implications must be carefully managed.

Regarding citizen participation, he believes motivation can be increased through monetary incentives (e.g., sharing a portion of fines for reporting idling trucks), gamification (leaderboards, points), or a civic contribution model. He emphasized that any strategy should balance motivation, engagement, and data quality appropriately.

For the future platform design, Dr. Mydlarz noted that if agencies plan to develop a next-generation participatory monitoring and reporting system, they should consider the following:

1. Define clear expectations for stakeholders: This is a foundational step that defines the technical implementation, the allowable error, and type of data that will be collected. The reason that data is needed, and how much error is acceptable should be defined in terms of the end user agency.

2. Coining hybrid sensing systems: One ideal system will be composed of the citizen-science data and the fixed-sensor network because both complement each other. Citizen engagement will deliver high–spatial-resolution data for first-hand observations at many locations and fixed sensors will deliver high–temporal-resolution data that continuously monitor for detailed investigations of local environment.The linking of citizen reported data to official data measured by sensors significantly strengthens the overall utility, credibility, and interpretability of the data.

Also this planning should take into account the opportunity to plan for a platform or application, which is able to host several environmental monitoring mechanisms, such as noise, air-quality, flood record, while being fully integrated to existing data systems available in the city to enhance efficiency and maximize adoption.

Dr. Norouzi provided a constructive critique on citizen science projects from the design and governance dimensions, stating that a fundamental pitfall in too many citizen science projects is a lack of community-driven ownership. Too often, the design of the project was top-down, “making for” people in the community instead of “making with” the community, which has the unintended consequence of side-lining tacit knowledge that exists at the local level and that, failing to match the real capacity and real interest of the community, might result in short lived engagement and participation. For this reason, he has advocated for more engaging and participatory design principles that deliberately lower the entry barriers to activities, tools, and interfaces, making them accessible regardless of a user's technical competence or prior exposure to science.

Regarding participating communities, Dr. Norouzi contended that it is important to articulate the value of a project and talk about that value to interested and direct users to keep support on the side of engineers, citizen science, and other community members. One specific point is to identify a target group especially with a distinct "pain point", that is a straightforward environmental concern as immediate concern to the impacted audience, and subsequently talk about how the project can bring value to end-users later on.It is essential to this audience-centred perspective because it creates a way to move the participation from the type of short-term incentivised participation, to long-term internally motivated participation.

Given his concerns on risk and his emphasis on the need to demonstrate feasibility, he strongly advocated piloting the work with proof-of-concept studies or pilots studies prior to full-scale implementation; the need for pilots especially in the form of realistic data collection setups (such as on smartphones) being particularly important for demonstrating feasibility, and of finding unknown issues with the technical and usability of the approach.

Second, Dr. Norouzi also identified data quality as a critical problem, connecting it to the preparation process. He proposed these first stages as not just proof-of-concepting engagement methods, but as ensuring that they also develop baseline recognition of the data variance present on consumer devices in order to develop appropriate calibration methods, and possibly to shape expectations of how the data could be used. Side-by-side with this was the need for privacy and trust. He claimed that rigorous ethical data governance, such as open communication about how the data was being used, how they were stored, and how data was made anonymous, would be necessary to engender public trust as a social license, and avoid the failure of the project.

Lastly, Dr. Norouzi proposed a generative framework as a conceptual extension of the topic. He suggested that citizen-science projects could adopt a multisensory design heuristic, organizing activities around documenting environmental phenomena through sight (visual pollution, skyglow), sound (noise), smell (airborne odors), touch (heat-island effects), and even taste (implications for water quality). This multisensory approach would democratize participation by enabling experience-based interfaces and, in theory, has the potential to produce well-rounded, multidimensional environmental datasets that support a more equitable form of citizen science.

7.2 Insights from public-facing experts

Interviews on how to authenticate, analyze, and operationalize citizen-generated environmental data with Dr. Hannah Eisler Burnett of the Science and Resilience Institute at Jamaica Bay (Sea Grant, FloodNet partner) and Dr. Lauren Smalls-Mantey of the NYC Department of Health and Mental Hygiene, provided insights into the advantages and disadvantages of utilizing smartphones and community involvement to augment conventional monitoring systems. Their insights overlapped and offered guidance for noise-validation work as well as future cross-hazard integration.

Dr. Smalls-Mantey emphasized that citizen research consistently increases coverage in both temporal and spatial terms, particularly in urban areas where fixed instruments are either few or expensive to set up and maintain. In low income areas, where environmental burdens are unevenly dispersed, public participation can reveal hyperlocal trends that formal systems miss, as demonstrated by her experience overseeing community-led infrastructure monitoring and heat adaptation programs. She did, however, also emphasize the serious methodological difficulties can occur when the general population gathers data without scientific training. High error margins or erratic patterns in the dataset can result from inconsistent measuring methods, variations in device hardware, and “everyday randomness” in human behavior. She clarified that "you'll see a lot of different strange things…people are weird" when community members monitor environmental variables without following tight rules, and that data can quickly become noisy or challenging to compare between users.

This highlights the need for integrated limitations, cues, or instructions that enhance environmental measurement consistency in the context of the data collected by the new NYC

Noise Monitor App. Dr. Smalls-Mantey recommended limiting measurements to certain locations, forcing users to contribute contextual photos, or using controlled measurement periods (e.g., requesting participants to record at the same time every day). By establishing more controlled sample circumstances, these tactics would assist NYC DEP and MOCEJ in lowering variability brought on by user behavior or environmental randomness. Given the wide difference in microphone sensitivity among phone models, production dates, and even cleanliness, she also emphasized the importance of baseline calibration measurements. According to her advice, NYC DEP and MOCEJ should include device-model metadata, fast baseline calibration procedures in the app, or backend post-processing adjustments in its future data validation framework. This aligns directly with findings from the literature review, discussed in Section 5.1, which similarly emphasize device calibration and model-specific correction as essential steps for improving the reliability of citizen-generated noise data.

Dr. Smalls-Mantey also shared ideas from her experience managing citizen science initiatives with and without participant pay as it relates to community involvement and retention. She pointed out that while remuneration greatly improves the quality and consistency of involvement, its viability depends on steady funding, which has historically constrained many city-run initiatives. Seniors and ecologically conscious groups, she added, are dependable supporters since they are more likely to stick to routines. This implies that before trying citywide scalability, NYC DEP and MOCEJ's future outreach strategy might start with driven, interested citizen groups. Dr. Smalls-Mantey's remarks about the unpredictability of public participation highlight the need to weigh user engagement tactics, such as gamification, notifications, or community groups, against the possibility of producing "garbage data" if incentives favor quantity over quality.

Dr. Burnett's perspective complements Dr. Smalls-Mantey's by situating citizen science within the broader ecosystem of environmental sensing and public data infrastructure. As a researcher involved in FloodNet and the MyCoast platform, Dr. Burnett described how community-generated flooding photos and stationary sensor readings can validate one another. Validation is a two-way process. Sensor data may validate citizen inputs, community data can validate sensor readings. "Under what conditions can citizen noise reports serve as evidence alongside sensor data, 311 complaints, or city inspections?" might be a question to address in addition to "How do we validate citizen noise data?" Dr. Burnett stressed that different validation techniques are needed for various environmental occurrences, such as floods, heat, air quality, or noise.

Dr. Burnett stressed that the validation technique depends on the purpose of the data. Mobile measurements would be more useful if NYC wants to evaluate noise exposure as people travel through the city. Timestamps, location accuracy, and consistent recording methods are especially important if the objective is to confirm site-specific noise breaches. Additionally, she advised choosing a single noise category as a pilot use case, such as construction noise or standardized emergency sirens, instead of validating all noise categories at once which would spread the project too thin. This concept is consistent with the more general use of scoped pilots in citizen science projects.

Dr. Burnett's insights on community engagement were especially rich and relevant to future expansion beyond noise. She made a distinction between three frequently used terms: community-engaged research, public engagement, and user experience. A tool's functionality and ability to address a real-world issue are guaranteed by the user experience. Building residents' awareness, interest, and trust is the main goal of public involvement. Community people are directly involved in the design, collection, and interpretation of the data in community-engaged research. She pointed out that FloodNet's effectiveness is organizational as well as technological; it prioritizes developing reciprocal connections over approaching engagement as public relations or data harvest, has a dedicated full-time engagement staff person, and does continuing community education.

A consistent theme emerges from both interviews: citizen science can meaningfully expand environmental monitoring capacity, but only when paired with thoughtful validation criteria, practical metadata collection, and genuine community engagement. This entails creating procedures for NYC's noise project that take into account user unreliability, device unpredictability, and the requirement for precise measuring procedures. These same ideas hold true for future environmental expansions, such as heat, flooding, and air quality, but the data types, calibration needs, and community objectives will change. The insights above support the idea that validation is a sociotechnical process requiring public trust, ongoing participation, and reciprocal value between agencies and communities rather than being just a technical activity.

7.3 Insights from private experts

Insights from the private sector highlight the operational, design, and engagement challenges associated with incorporating citizen-generated environmental data into public-facing systems. This section draws on interviews with Nate Rauh-Bieri, Head of Projects at JustAir, and Dr. Eben Cross, Chief Science Officer at QuantAQ, both of whom have extensive experience developing and deploying air quality sensor networks.

Both experts emphasized that citizen-generated data is most valuable when it is part of a cooperative structure involving community organizations, researchers, and government partners. Rather than functioning as an isolated dataset, community-collected information tends to be most impactful when it answers questions relevant to local concerns or fills spatial and temporal gaps that formal monitoring systems cannot cover. They noted that citizen sensing is particularly useful for highly localized issues, such as block-level hotspots or recurring disturbances concentrated in specific neighborhoods. In contrast, for broad, uniform environmental events, such as regional wildfire smoke, low-cost networks add little new insight because all sensors tend to show the same pattern. This distinction suggests that, for noise, citizen participation would be most meaningful when tracking neighborhood-specific conditions rather than attempting to characterize citywide trends.

Another shared insight was that data quality cannot be taken for granted. The experts stressed that low-cost or citizen-driven data becomes difficult to use when devices lack calibration, when installations occur in inappropriate locations, or when machine-learning models produce opaque results without clear interpretability. Agencies often hesitate to rely on such data because they cannot verify how it was collected or processed. They described more rigorous practices used in their own projects, which include baseline checks built into the device, additional quality assurance applied when data reaches the cloud, direct calibration of each unit before it is deployed, and periodic comparison with regulatory-grade instruments. They advised that data validation needs to be clearly aligned with the standards and expectations of the public agency.

For noise, this means explicitly defining what accuracy thresholds are acceptable relative to the NYC Noise Code and testing smartphone recordings against professional acoustic equipment to understand their limitations and potential use cases.

Both interviewees also pointed out that even technically sound systems fail when user experience is neglected. Engagement declines rapidly when interfaces are confusing, onboarding instructions are overwhelming, or visualizations are too technical for the general public. They emphasized the importance of clear and accessible designs, particularly for older adults or residents with limited digital literacy. Small interaction prompts, such as asking users to add context, annotate unusual conditions, or confirm readings, can significantly improve the quality and interpretability of the collected data. They also noted that public-sector tools must compete with other apps for users' attention; if a tool does not immediately communicate usefulness or feel intuitive, participation tends to drop quickly.

Perhaps most importantly, both experts warned that even well-designed and scientifically validated systems struggle to make an impact when they are not fully integrated into existing government workflows. There have been cases in air quality projects providing accurate community-generated data failed to influence decisions because inspectors continued relying on older procedures or because the data did not align with established complaint pathways. This has direct relevance for NYC's noise context: even if citizen data become technically reliable, they will only contribute meaningfully if agencies adjust their protocols to incorporate and act on these new information sources.

7.4 Interview synthesis

A recurring theme from all of the interviews was that citizen-generated environmental data is only useful when combined with precise guidelines, careful validation procedures, and appropriate community involvement. Academic experts emphasized the importance of mixed systems that combine citizen sensing with fixed monitoring networks, while public-facing experts highlighted practical constraints such as device variability, user inconsistency, funding limitations, and the sociotechnical nature of public trust. Private-sector experts reinforced that rigorous calibration, metadata collection, and workflow integration are essential for making citizen data actionable within government processes. Together, these perspectives suggest that the new NYC Noise Monitor App, and any future multi-hazard platform, must balance usability with scientific rigor, designing tools that both produce data reliable enough for policy use and empower residents. In the end, the interviews demonstrate that citizen science can be successful because citizens and agencies collaborate to produce a more comprehensive and equitable picture of environmental conditions rather than because the public replaces formal processes.

8. Review of new NYC Noise Monitor app

The NYC Noise Monitor mobile app, launched by DEP in November, 2025, enables residents to measure ambient noise levels using their smartphones and submit anonymized reports containing decibel readings, noise-type classifications, and location information. Users are required to select a noise category, such as traffic, construction, music, alarms, or other sources, when submitting each measurement, allowing the platform to organize observations by noise type and support high-level pattern recognition. Designed to expand the spatial coverage of noise data beyond the limitations of fixed sensors and 311 complaints, the app represents an important step toward participatory noise monitoring and reporting in New York City.

A review of the app's structure reveals several strengths. Its simple interface lowers the barrier for participation and allows individuals to record measurements quickly without technical expertise. By requiring users to classify noise sources, the app provides metadata that can help contextualize observations and differentiate patterns of urban noise. Mapping tools further support the identification of recurring hotspots and temporal trends, aligning with DEP's goal of improving understanding of citywide noise conditions and guiding targeted enforcement efforts.

Despite these advantages, however, the app presents notable limitations that affect data quality, user trust, and long-term engagement. The absence of more detailed contextual metadata, such as phone placement, environmental conditions, or motion state, limits interpretability and complicates comparisons across contributors. It also remains unclear how app-based measurements will be incorporated into DEP's enforcement workflows.

User experience concerns further challenge sustained engagement. Some feedback indicates that users find the measurement process unclear, including uncertainty about what the app is measuring and difficulties resetting or revising submissions. Reviews also note issues such as accidental report submissions, the absence of a confirmation step, and the inability to retract or delete erroneous entries. More importantly, the app provides limited personal benefit or feedback, making it difficult for individuals to understand how their contributions lead to meaningful outcomes and the willingness to keep using the app.

Overall, the new app offers a promising foundation for participatory environmental monitoring. Its long-term effectiveness, however, will depend on improvements to data validation, metadata capture, user experience design, and the integration of citizen-generated information into trusted city agency operations.

9. Future trends

9.1 Future vision

Looking ahead, citizen-generated data can play a much larger role in how cities understand and respond to environmental issues. As traditional sensor networks remain limited by cost and spatial coverage, public participation offers a scalable way to fill gaps in real-time monitoring. Increasing interest in community science, both in New York City and globally, suggests that residents can become more closely involved in documenting environmental conditions and contributing to early-stage problem identification (Cappa et al., 2022; Kontokosta et al., 2016).

9.2 Project development

As the new NYC Noise Monitor App becomes operational, our team expects community engagement to become an important next challenge. While citizen scientists can contribute a substantial portion of the data, some deployment of fixed sensors across the city will still be necessary to verify and cross-check community-generated measurements.

A practical first step in addressing this challenge is to engage interest groups and neighborhood organizations that are already active around environmental or quality-of-life issues. Small community events or hands-on workshops can also help raise awareness and encourage residents to participate. Because the new NYC Noise Monitor App relies heavily on citizen science, expanding the user base, both in size and diversity, will directly strengthen the quality and usefulness of the data. Lessons from FloodNet show that strong partnerships and visible community involvement are critical for sustaining participation over time.

9.3 Long-term goal

Further into the future, a trend of numerous cities towards integrated environmental platforms is developing which pulls together multiple hazard types together in a single public-facing display. Rather than servicing multiple dashboards, agencies are exploring more shared infrastructure which would lend itself to, for example, easier cross-hazard comparison, better real-time visualization, and more accessible communication with residents and policymakers. The broadening of open data policies and of coordination among agencies are increasingly making this sort of consistent data repository possible than it was in the past. Collectively, these trends represent a sea-change that says that, in the future, our environmental governance is likely to come from places that use data from a range of sources and which is accessible and understandable to the public.

10. Conclusion

This capstone project examines how New York City can leverage citizen-generated noise data to support environmental monitoring and decision-making. Using a mixed-methods approach, the project demonstrates that smartphone-based citizen-generated data has the potential to complement existing noise-monitoring infrastructure. At the same time, the findings highlight several critical challenges, including data validation, system integration, and sustained community engagement, that must be addressed before such data can be reliably incorporated into operational systems.

The findings show that citizen-reported noise measurements can help fill critical spatial and temporal gaps that persist in 311 complaints and fixed sensor networks. This potential hinges upon an organized data validation strategy, thoughtful user interface design, and agency workflow. Important challenges, including heterogeneity of devices, calibration, sampling biases, privacy issues, and the ability to engage in enforcement, must be addressed effectively before citizen science would be available for use in trusted city agency operations.

The analysis also underscores the role of contextual metadata, device aware calibrations, contributor reputation, and cross-validation using reference sensors, as well as the role of communication to users about the intent and limitations of their data contribution. In addition to the technical concerns, interviewees revealed that lasting citizen science must be based on trusted partnerships with community groups, careful incentive design, and an accessible, fun user experience.

Looking ahead, the launch of NYC DEP's new Noise Monitor App represents a major milestone toward participatory environmental monitoring. In the future, such an approach could support a unified public dashboard that integrates multiple environmental stressors, including noise, air quality, flooding, and heat, thereby enabling more transparent, responsive, and data-driven urban governance.

11. References

Bello, J. P., Silva, C., Nov, O., Dubois, G., Arora, A., Salamon, J., … Mydlarz, C. (2018). SONYC: A system for the monitoring, analysis and mitigation of urban noise pollution. Communications of the ACM, 61(2), 68–77.

Cappa, F., Rosso, F., & Hayes, A. (2022). Citizens and cities: Leveraging citizen science and big data. Business Strategy and the Environment, 31(1), 1–15.

Kontokosta, C. E., Hong, B., & Johnson, N. (2016). The quantified community at Red Hook: Urban sensing and citizen science in low-income neighborhoods. Buildings, 6(4), 43.

Picaut, J., Boumchich, A., Bocher, E., Fortin, N., Petit, G., & Aumond, P. (2021). A Smartphone-Based Crowd-Sourced Database for Environmental Noise Assessment. International Journal of Environmental Research and Public Health, 18(15), 7777.

Zhu, J., Li, J., Liu, L., & Tham, C. K. (2015). iCal: Intervention-free calibration for measuring noise with smartphones. In Proceedings of the 2015 IEEE 21st International Conference on Parallel and Distributed Systems (ICPADS) (pp. 244–251). IEEE..

12. Appendix

Appendix A: Interview with Dr. Charlie Mydlarz

Q1: What is your general view on citizen science data, and how do you think it helps cities monitor environmental issues?

A: My Ph.D. research actually involved using citizen scientists to capture noise data via smartphones. I have always been a strong proponent of citizen science. The main reason to use it is when you need data from areas that are difficult to access with fixed sensors. Generally, sensors provide better temporal resolution since they continuously monitor once installed, but they lack spatial coverage. Citizen scientists, on the other hand, can provide broad spatial coverage because they are mobile and can reach diverse locations.

However, there are trade-offs. Human involvement introduces potential errors in the measurement process. Smartphones vary in condition and calibration, so the data quality can be inconsistent. While averaging across large datasets can help compensate for these inconsistencies, variability remains an inherent challenge.

Q2: What are the major challenges you have encountered in citizen science projects?

A: Recruitment is one of the biggest challenges. The success of a citizen science project depends on having a large and active user base. Because of potential data quality issues, projects rely heavily on having a high volume of data, which means motivating many participants.

There must be a clear incentive for participation. Some cities use monetary rewards. For instance, there are programs that pay participants for reporting idling trucks, even giving them a share of the enforcement ticket revenue. Others use gamification techniques, such as leaderboards, to engage users. However, these methods can be a double-edged sword—they may encourage more data submission but also increase the risk of low-quality or false data.

Ultimately, successful projects must find a balance between recruitment, motivation, incentive design, and data quality. There is extensive literature discussing these trade-offs in citizen science.

Q3: What are the main technical issues when using smartphones to record noise data?
A: Smartphones can provide reasonable approximations of noise levels, but they are not designed as precision instruments. Calibration is a significant issue—it requires specialized equipment and effort, which most users are unwilling or unable to perform. The condition of each phone also affects data accuracy. For example, microphones may be clogged with dust, resulting in lower sensitivity.

Different devices of the same model can produce varying measurements. Therefore, while phones can indicate general loudness levels (e.g., 80–100 decibels), they cannot reliably provide exact decibel readings within a narrow margin of error.

In enforcement contexts, such as city noise monitoring, it’s essential to clarify with stakeholders (e.g., the Department of Environmental Protection) what level of accuracy they require and how they intend to use the data. If the goal is to identify general noise hotspots rather than issue fines, smartphone data can be entirely valid and useful.

Q4: Can calibration or post-processing improve data accuracy?
A: Many noise-monitoring apps include built-in compensation algorithms that adjust for device models, which improves accuracy to a degree. However, these adjustments cannot account for the physical condition of each individual phone. True calibration requires testing each device against a known noise source in controlled conditions—something impractical for large-scale citizen science projects.

Therefore, developers must determine with stakeholders what margin of inaccuracy is acceptable for their specific use case. The focus should be on how much variability can be tolerated rather than achieving perfect precision.

Q5: Beyond noise, can citizen science methods be extended to other environmental issues?
A: Absolutely. Leveraging mobile devices for data collection is a powerful and scalable approach. The key is to align expectations between the developers and the end users—in this

case, the city and its agencies. Understanding how the data will be used and the required level of reliability is crucial.

Citizen science data doesn't always need to complement fixed sensors, but it can significantly enhance existing data systems. For instance, in New York City, noise complaints typically originate from 311 reports. These complaints trigger investigations and inspections. By adding smartphone-based noise measurements and audio samples, inspectors can access richer contextual information before deploying field teams. This improves efficiency and helps target enforcement efforts more effectively.

Q6: What privacy concerns should be considered when collecting noise or audio data?
A: If an app records only decibel levels, privacy concerns are minimal since decibel values cannot be reverse-engineered into speech. However, recording audio introduces privacy risks, as it may capture identifiable voices or sensitive information.
To address this, the app should include a clear disclaimer or consent form stating that users will not record others without permission. Legal review is necessary to ensure compliance. Based on my experience, visual signage or disclaimers are effective measures when devices or sensors are placed in shared or public spaces.

Q7: How should location data be handled for environmental monitoring apps?
A: I recommend collecting precise latitude and longitude data from the device. With this, you can later derive contextual information such as nearby intersections or addresses. However, GPS accuracy can fluctuate, especially indoors or in dense urban areas like Manhattan.

A practical approach is to initiate GPS tracking as soon as the app opens to allow time for the signal to stabilize before taking measurements. The app can also record the accuracy value of each reading for transparency and quality control.

Q8: If you were to design the next generation of participatory environmental monitoring platforms, what would your top priorities be?

A: First, I would prioritize establishing clear expectations with stakeholders, as this defines all technical and methodological decisions—from data accuracy standards to collection methods. Second, I would focus on integrating citizen science data with other data sources.

The most effective systems are hybrids, combining fixed sensors for temporal coverage with citizen scientists for spatial coverage. For example, a flood-monitoring platform could allow users to submit geotagged photos alongside sensor data. Such a system would not only provide quantitative measurements but also rich visual context, enhancing the overall understanding of environmental events.

Q9: Any final thoughts?
A: Citizen science is a valuable and evolving tool for urban environmental monitoring. The key to success lies in balancing data quality, participant motivation, and stakeholder expectations. With thoughtful design and clear communication, citizen-generated data can meaningfully support city management and environmental decision-making.

Appendix B: Interview with Dr. Hannah Burnett

Q: What are your thoughts on how citizen-science-based data—such as crowdsourced environmental data—can complement traditional monitoring methods like fixed sensors? From my understanding, FloodNet uses sensors around the city to detect flood occurrences, while MyCoast relies on community members uploading photos of flooding. How do these two types of data—sensor-based and community-based—work together or complement each other?

A: That's a great question. They don't quite work together yet, but they address similar challenges.

MyCoast started in 2018 and has gathered a large collection of photo data contributed by residents documenting flooding events. Since around 2021 or 2022, this data has been hosted by NYC MOS. On the MyCoast website, you can see that the uploaded photos are linked to traditional environmental datasets—such as tide gauge readings and recent precipitation data.

When users upload photos, the metadata, including geolocation and timestamp, is automatically captured. This allows MyCoast to link community reports to conventional sensor-based data.
As for FloodNet, the project was initially launched in part as an effort to validate citizen-generated flood data. The idea was to empower residents to build tools that would help them advocate for flood-adaptive measures within their communities.

So in some sense, FloodNet's purpose is almost the inverse of what you're asking about—it was designed to help validate community-collected data rather than the other way around.
However, there's a reciprocal relationship between these data types. Citizen-generated data can help validate sensor data—for example, confirming whether sensors are accurately detecting floods—and vice versa. The interplay really depends on the objective of the project and the environmental phenomenon being monitored.

Flooding, for instance, is very different from something like heat. But community science projects can significantly enrich traditional datasets by providing context, lived experience, and human-scale information.

Let me give you an example: the Climate, Heat, and Air Quality Mapping Project by the New York Environmental Justice Alliance, also known as the CHAMP EJ Report. In that study, community members helped monitor temperature and air quality throughout a very hot day.

What they found was fascinating—people's exposure to pollutants, such as particulates from traffic, was much higher than what stationary sensors indicated. When sensors are fixed and record averaged data, they can't account for how people move through the city—whether by walking, biking, or commuting—which greatly affects their actual exposure.

That's a powerful example of how citizen and community data can enhance and contextualize traditional monitoring approaches.

Appendix C: Interview with Dr. Hamidreza Norouzi

Q1: What do you think are the main challenges in ensuring the accuracy and reliability of citizen science-supplied data (including crowd-sourced data)?

A:

- Should give ownership to the community
- Design has often been from top (engineer) to bottom (community)
- Design should be inclusive (not to differentiate users based on their knowledge level)
- Community engagement is hard

Q2: In your view, what are effective ways to encourage sustained citizen participation in environmental monitoring?

A:

- Find the painpoint (helping their community)
- Find target audience
- The project should be beneficial to them

Q3: What else do you think is important for citizen science data?

A:

- Calibration: Do some trial before the project started (e.g. actually collect data on phone)
- Privacy issue: Building the trust and Protect the data

Q4: If you were to design the next generation of participatory environmental monitoring platforms, what would your top priorities be?

A: Everything of the five senses can be considered to start a new project.

Appendix D: Interview with Dr. Lauren Smalls-Mantey

Q1: What are your general views on using citizen science to help monitor environmental issues like noise or air pollution?
A: Citizen science is a great tool. It adds an extra layer of data—both quantitative and qualitative. Citizens can help collect measurements or provide contextual feedback, which strengthens environmental monitoring in the city.

Q2: Can you describe your past experience with citizen science or community-based monitoring programs?
A:

- During my dissertation, my lab created a citizen scientist program for New York City's green infrastructure. This was before DEP's wide-scale rollout of bioswales and rain gardens.
- We trained community members on green infrastructure, how to maintain sites, and how to collect maintenance data.
- The program ran for about three years and produced information on how often these sites needed cleaning.
- At NYC DOHMH, I worked on Cool Neighborhoods NYC, which deployed ~500 air temperature sensors. Community involvement was limited, but we ran a "Be a Buddy" peer-to-peer heat–risk system.
- We also ran an indoor temperature sensing program where volunteers had temperature sensors in their homes and could view their data on a dashboard. It had mixed results.
- Throughout my career I've also used community sensor networks like Weather Underground, but always inspect data validity before using it.

Q3: What are the main challenges in ensuring the accuracy of citizen-generated noise data from smartphones?
A: Consistency is the biggest challenge. Everyday people don't always follow protocols, so data quality is more variable. Phone mics also differ widely by model, age, manufacturing date, and physical condition. Dirty microphones (lint, grease) affect readings. These sources of variation make noise measurement particularly difficult.

Q4: Given these limitations, what methods could improve reliability when using smartphones for noise measurement?

A:

- Don't attempt continuous measurement. Instead, ask people to record at specific times or within defined windows.
- Encourage people to take measurements in the same place each time, or collect a photo of where they are.
- Consider offering predefined location options.
- Build in redundant measurements: ask participants to record twice within about five minutes. This catches errors or anomalies.
- Introduce a baseline calibration step in the app before each recording, since microphones vary so much.
- Ideally, compare some smartphone readings with co-located calibrated noise sensors to understand typical error margins.

Q5: What recruitment strategies make sense for a citizen-science noise project?

A: Start with people already bothered by or sensitive to noise—environmental groups, neighborhood associations, or residents who complain about noise. They are motivated and more likely to participate consistently.

Q6: What factors affect the success or scalability of community engagement efforts?

A:

- Funding is critical. Many city programs fail to continue simply because they are not refunded, even if they're effective.
- Compensation significantly improves participation and retention. Programs with compensation perform better than those without.
- Consider different age groups—seniors, for example, can be a very steady cohort. With support, they can learn to use the app and are reliable because they have stable routines.

Q7: What are your thoughts on using gamification to improve engagement?
A: Gamification itself isn't necessarily harmful, but campaigns that shared links widely online have run into bot issues. To avoid illegitimate participation, recruitment should be done in person or in controlled environments. Don't publish a public link too broadly.

Q8: Beyond calibration and sensors, what other challenges should we expect when working with citizen-supplied data?
A: People are unpredictable. You'll see strange patterns and inconsistencies in responses—this happens in surveys or sensor-based projects. To manage this:

- Build redundancy (e.g., duplicate measurements).
- Use intake and exit surveys to understand participant behavior.
- Expect a lot of "noise" in the data—not just acoustic noise, but behavioral noise.

Q9: How should we think about validation for the noise pilot?
A: True validation is hard without having sensors in the area. You likely need to combine smartphone data with at least a few calibrated noise sensors. If DEP already has sensors, explore those; otherwise, try to obtain a couple of small sensors so participants can test phones side-by-side. That comparison helps determine the maximum acceptable error for your pilot.

Q10: Do you have any final thoughts for us as we design our study?
A: Focus on defining measurement windows that make sense for your semester timeline. Consider what times of day generate the most useful data. Always assume the data will be messy and design methods—redundancy, surveys, calibration—to manage that reality.

Appendix E: Interview with Nate Rauh-Bieri

Q1: To begin, what is your general view on using citizen science to support city-level environmental monitoring?
A: For example, when discussing disparities between monitors, community input helps us understand whether unexpected readings reflect real environmental conditions or require further investigation.

Q2: What do you see as the main challenges in ensuring accuracy and reliability of crowdsourced environmental data?
A: Our approach emphasizes validation: installers follow standard procedures, we perform regular maintenance, and we have system-level checks to flag anomalies. If a reading seems suspicious, we can temporarily remove a monitor from the dashboard or send clarifying messages to users.

Q3: How can a city make a platform accessible and appealing to a wide range of residents, and encourage strong citizen participation?
A: Communication design is critical. Even expert researchers struggle to translate technical data into accessible visuals and messages. Good UI/UX is its own form of communication.

Another opportunity is building interactive or gamified feedback loops. People naturally engage with light gamification.

We originally built our platform through community groups focused on environmental justice, which gave us a strong foundation. However, one area we wish we had invested in earlier is designing for users who are not digitally fluent—such as older adults.

Elderly residents, for example, care deeply about environmental conditions and would benefit from simplified sign-up flows and more intuitive interfaces. They may also be a high-value target segment for a noise-monitoring tool.

Q4: In your experience, what are the biggest risks or challenges cities face when leveraging citizen science? This can relate to noise monitoring or citizen science more broadly.
A: One significant challenge is the attention economy. People are accustomed to billion-dollar apps with flawless interfaces. If an app lacks polish, users may not have the patience to engage with it, especially when notifications compete with every other app on their phone.

Another challenge is balancing lived experience with quantitative data. Lived experience is invaluable—communities often know what is happening long before the data reflects it. But policymakers often rely on data to legitimize action. Platforms must navigate how to integrate and validate both forms of knowledge respectfully and effectively.

Additionally, screening the wide variety of community input can be difficult. Designing a structured way to collect, filter, and interpret feedback will be essential for any citizen science initiative.

Appendix F: Interview with Dr. Eben Cross

Q1: What models have you observed in community or research deployments of your sensors?
A: Our most successful deployments form a coalition among a community-based organization, researcher, and local government partner.

When these three work together, combining community knowledge, scientific rigor, and regulatory authority, the project becomes impactful. Our sensors sit at the center of this triangle, helping answer local questions about what is happening, why it happens, when, and for how long. This information supports decisions at policy, public-health, and individual levels.

Q2: How are sensors installed, and what safety considerations are involved?
A: We design the system so non-experts can install sensors securely on existing vertical infrastructure. We provide clear installation guides—video and written. Customers install the sensors; we do not.

Mechanically, the key safety concern is preventing heavy hardware from falling. As long as the installation is correct, the risk is minimal. We avoid installation liability while ensuring the equipment is difficult to install incorrectly.

Q3: Do you perform data validation? How?
A: Yes. We apply QA/QC at several stages:

- On-device: second-level QA/QC.
- Cloud: additional QA/QC when data is aggregated into one-minute intervals.
- Pre-deployment: all sensors are calibrated in our in-house calibration chamber.

The largest challenge in lowering monitoring costs by 100× is ensuring inexpensive sensors still produce data comparable to state-run, high-grade equipment. Maintaining known data-quality objectives is central to our work—as atmospheric scientists, we value accurate and credible measurements.

Q4: What metadata do you collect, and how do you maintain transparency?
A: We are fully transparent. Customers can access:

- Raw data: direct voltage signals from sensor components.
- Final processed data: pollution concentrations.
- The full model: including assumptions and transformation steps.

This allows users to trace exactly how a final value was derived and avoid "modeling a model," which happens when companies hide raw data behind opaque machine-learning layers. Noise may be less complex than air quality, but accuracy still matters because reliable decibel measurements determine whether the government may use the data for regulatory purposes.

Q5: Which validation methods for air pollution might transfer to noise monitoring?
A: The key analog is determining whether low-cost microphones are "good enough." Professional microphones cost around $20,000; smartphone microphones cost about $2.

The question becomes: How do you define "good enough"? Can a smartphone microphone meet the standard set by NYC DEP noise ordinances?

If validated empirically—by comparing a citizen device to professional-grade equipment—the potential for large-scale, citywide noise monitoring becomes compelling.

The challenge, as with air quality, is connecting this data to a meaningful regulatory workflow.