



**NOVEMBER 2024**

# **A FRAMEWORK FOR CITIZEN- DRIVEN DATA SHARING IN URBAN POLICY-MAKING**

**KHAHLIL A. LOUISY**



NOVEMBER 2024

# A FRAMEWORK FOR CITIZEN- DRIVEN DATA SHARING IN URBAN POLICY-MAKING

KHAHLIL A. LOUISY

# About Data-Smart City Solutions

**Data-Smart City Solutions** at the Bloomberg Center for Cities at Harvard University is working to catalyze the adoption of data projects on the local government level by serving as a central resource for city leaders. We highlight best practices, top innovators, and promising case studies while also connecting leading industry, academic, and government officials. Our research focuses on the intersection of government and data and explores innovations in open data, predictive analytics, and civic engagement technology. We seek to discover and preemptively address civic problems by integrating cross-agency data with community data.

## About the Author

**Khahlil Louisy** is a contributing author for the Data-Smart City Solutions program at The Bloomberg Center for Cities at Harvard University and a former Technology & Human Rights Fellow at the Carr Center for Human Rights Policy at the Harvard Kennedy School. Louisy is an applied economist focused on issues of public and global health, economic development, and technology and innovation. His work centers on the development and application of technologies for public purpose, while researching their implications for issues of inequality, health outcomes, and human rights. He is the former Head of Global Implementation at PathCheck Foundation - an organization founded at the Massachusetts Institute of Technology (MIT) to develop novel technologies in response to health emergencies. Louisy currently serves as President of the Institute for Technology and Global Health and Co-Head of AI and Technology for Public Health -Outbreaks, within the joint World Health Organization (WHO) and International Telecommunications Union (ITU) initiative on Artificial Intelligence for Health. His work has spanned several countries globally and he remains committed to issues of equality, equity, and global poverty.

# Acknowledgements

This paper was edited by Betsy Gardner, senior editor for Data-Smart City Solutions at the Bloomberg Center for Cities at Harvard University.

This paper is copyrighted by the author(s). It cannot be reproduced or reused without permission. This paper is available to the public at [datasmart.hks.harvard.edu](http://datasmart.hks.harvard.edu) free of charge.

 DATA-SMART  
CITY SOLUTIONS

**A PUBLICATION OF**

**Data-Smart City Solutions**

79 John F. Kennedy Street Cambridge, MA 02138

[datasmart.hks.harvard.edu](http://datasmart.hks.harvard.edu)



## Introduction

Today's urban landscapes are rapidly evolving, and data has become a crucial asset for cities and policymakers to address complex challenges, from health and public safety to traffic congestion and climate change. Yet, as cities adopt data-driven approaches, there is a growing recognition of the following two points. First, effective policy making relies not only on data and technologies, but also on the active participation of citizens. And second, datasets available to policymakers, primarily from public sector agencies, may be incomplete and do not provide a comprehensive overview of existing challenges. For example, data from a city's health department may not accurately reflect the number of residents affected by cardiometabolic diseases or respiratory illnesses due to environmental factors. Nor would data from a city's Department of the Environment alone truly reflect the hyper-local or street level air pollution conditions in neighborhoods. Citizen-driven data sharing, where residents contribute data from their personal digital devices like smartphones and watches, fitness trackers, and home sensors which tracks changes to environmental conditions in and around the home, provide a powerful and community-centered approach to designing responsive and transparent urban policies that addresses the needs of communities. Such initiatives help cities gather hyper-local and real time data to establish ground truths and deliver policies and interventions that are better informed and can be tailored to reflect the everyday realities of their diverse populations.

However, and unsurprisingly, establishing a citizen-driven data-sharing initiative is no simple task. To be truly effective, cities must first navigate a landscape of digital divides, privacy, and public trust concerns. Then there are the issues of storage and security, who has access, transparency and governance, community participation, and the nature of partnerships necessary to drive agility and efficiency that are commonly lacking in public sector undertakings.

This framework, rooted in voluntary participation and with a consent-based approach outlines the key elements of a successful citizen-driven data sharing initiative, derived from insights from established projects across several international implementations, published academic literature, and in alignment with existing data regulations. It addresses the issues of citizen participation, engagement, and autonomy, data security and usage policies, and cross-sector partnerships, all of which must be prioritized from the planning phase of the project.

## Framework for Data Sharing

### 1. Follow A Voluntary Participation and Consent-Based Model

The cornerstone of this framework is the voluntary participation of residents and their complete autonomy, including the freedom to choose the specific types of data they wish to share, such as air quality and particulate matter levels from their home monitors or heart rate and daily step counts from smartwatches. This is especially important if location data is included, such as latitude and longitude coordinates. Consent must be direct and clear, allowing users to decide precisely what data they're comfortable sharing and with whom, for example, city officials, universities and researchers, or the general public. Amsterdam's [Smart City](#) initiative uses simple and clear forms with explanations about data use and privacy protections, while Helsinki's [MyData](#) initiative allows residents to control data sharing settings and to view how their data is used.

### 2. Design Robust and Secure Data Anonymization and Aggregation Protocols

It is imperative that individuals' privacy is protected. All collected data should be anonymized and aggregated to a resolution that can still be useful, before it is analyzed or shared with policymakers. For example, data can be aggregated at the neighborhood or zip code level, rather than the household level. Aggregation across neighborhoods, zip codes, or zones within the city can help avoid the identification of specific individuals or households through deanonymization, while still offering valuable insights into localized trends. Where and how this data is stored is a critical factor to consider.

The secure anonymization and aggregation could be done via techniques like data masking and pseudonymization, differential privacy, and encryption, combined with strict governance protocols. Data masking removes personally identifiable information (PII) from datasets while pseudonymization replaces the PII with pseudonyms or tokens. This maintains the viability of the data for analyses but mitigates any potential for linking to specific individuals or households. Barcelona's [Sentilo Platform](#) utilizes this technique by gathering data from sensors across the city and removing any direct identifiers, instead focusing on aggregate information such as traffic, pollution levels, and temperature.

Differential privacy adds to the dataset in the form of "noise," by making small and random alterations to data points to protect individual privacy, while preserving trends in the data. This method minimizes the ability to reconstruct datasets or traceback to individuals, even if multiple datasets or sources are combined. Chicago's [Array of Things](#) - an urban sensing project - uses differential privacy to protect the data collected from environmental sensors across the city. This ensures that PII remains protected while insights into climate conditions, urban health, and infrastructure needs are uncovered.

Encrypting data, both while in transit and at rest, ensures that only authorized users may access and analyze the data, because encrypted data is unreadable to any entity attempting to breach datasets without authorization. San Francisco uses end-to-end encryption across its shared data platforms at each stage of transfer and storage.

Implementing strict governance and access controls will ensure that only authorized personnel can view and handle sensitive data and under what conditions. The Mayor's Office of Data Analytics (MODA) in New York City has implemented strict data governance measures that ensures that all personal identifiers have been removed prior to any analyses being performed and sufficiently aggregated prior to sharing with public policy teams.

In addition to each of these approaches, aggregation of data across sources and sectors is typically accompanied by the major challenge of interoperability. This is because the diversity of devices and data formats used within cities can lead to compatibility issues, making it difficult to aggregate and analyze data from various sources.

A solution to this problem is for cities to collaborate with device manufacturers and data management experts to develop and establish standard formats and protocols that drive interoperability. Barcelona's [CityOS initiative](#) provides a single platform for various data streams, enhancing data integration and usability for policymaking.

### 3. Build User-Friendly Platforms and Adopt Transparent Data Usage Policies

Platforms that are intuitive to use, where residents can manage their data-sharing preferences, see how their data is being used, and understand its impact on public policy is especially important for driving takeup. Transparency in data usage policies is critical for maintaining trust and ensuring citizens feel secure about participating.

Cities should adopt platforms that are visually appealing and simple, are accessible across devices (mobile phones, tablets, and laptops), and responsive. Incorporating easy-to-navigate elements like menus, the ability to easily choose and update data sharing preferences including for specific projects, and visualization tools will help residents to better understand and control their data contributions.

The language used on these platforms should also be clear and straightforward. Jargon-free language will make the information more accessible to residents and clear explanations about what data is collected, how it is used, and who has access to it will help residents feel both informed and empowered. Amsterdam's [Smart City Initiative](#) uses language that is transparent and plain on its data portals. The city explains the purpose for collecting each data type and how the data will be used. For example, noise levels data for urban planning and air quality for public health policies. Transparent data usage policies should be published openly on city websites that details exactly how the data is collected, processed, stored, and shared. Helsinki's MyData initiative is a great example of how cities can adopt clear data usage policies. The initiative promotes

transparent personal data usage across all its smart city projects and the policy is publicly accessible, which details governance practices and the specific applications of collected data. This approach has fostered trust among residents and driven take up rates and participation in data-sharing initiatives.

Regular publishing of transparency reports outlining usage, findings, and policy outcomes fosters accountability and informs residents on how their data is being utilized. New York City's MODA publishes [annual reports](#) that details data collected across various city services and their applications to public policies. Where transparency reports may not reach some citizens, hosting regular education information sessions and workshops can provide an avenue for residents to ask questions or get concerns addressed.

#### 4. Establish Cross-Sector Partnerships for Data Collection and Analysis

Collaborations between municipal governments, private sector entities like tech companies, and research institutions are crucial and can enhance data quality and depth. These collaborations increase access to specialized expertise and technologies, improve resource efficiency and cost sharing, accelerate innovation and knowledge exchange, improve agility in addressing complex urban challenges, and help foster global learning and benchmarking.

Cities and their personnel often lack both the specialized knowledge and advanced tools that private sector companies and top tier research institutions have. Cross-sector partnerships bring in additional experts which enhances the cities' ability to implement, organize, and analyze complex data.

Data collection and analysis themselves are resource intensive tasks, requiring both significant financial and human investment. Partnerships allow cities to share costs and resources, making large-scale data projects more feasible and enable cities to undertake more ambitious initiatives without the burden of budget constraints.

Partnerships also foster collaborative environments where ideas and methodologies can be shared across sectors. In particular, academic institutions can contribute through robust research and pilot studies that test innovative solutions to complex challenges. Cross-sector collaborations also accelerate innovation, enabling cities to adopt and adapt new strategies and technologies far quicker than if they did so in isolation.

Additionally, different sectors bring a range of diverse datasets and sources, insights, identification, and collection methods to the table, leading to richer datasets that offer a more comprehensive view of urban challenges. Further, we have seen that urban challenges like traffic congestion and public health are complex in nature and require multifaceted approaches. These partnerships allow cities to pull diverse perspectives and resources, making them more agile in responding to challenges as they evolve.

Involving a range of organizations, especially non-profits and academic partners, elevates trust within the community and helps bridge the gap between city officials and residents. A [survey on the adoption of digital vaccination credentials in the US](#) found that while the majority of the public (66.5%) supported its use, trust in the federal government and big tech companies were low. Support for local health departments and universities were highest for implementing and managing the data, while support for the development of the digital tool by the big tech companies was highest. These findings suggest that citizens generally prefer when institutions work together under different roles and responsibilities.

Cross-sector partnerships help cities gain insight into global best practices and set benchmarks for success. Cities are better able to adopt tested strategies from other cities and sectors and measure their progress, driving continuous improvement and positioning them as global leaders in smart city innovation.

## 5. Implement Feedback Loops for Resident Engagement

As trust is the cornerstone of this framework, strategies that build and maintain resident trust are vital. Implementing feedback loops in data-sharing initiatives is an essential element for building trust, sustaining engagement, and demonstrating the tangible benefits of shared data on urban policies.

Publicly accessible dashboards and data insights with live or regularly updated data should show how data trends influence policies. Producing reports and newsletters with visual explanations, like graphs contrasting before and after data-driven interventions helps residents understand the benefits of sharing their data and why decisions were made. This also helps to hold cities accountable if programs that were implemented are not driving any meaningful change.

Hosting regular community feedback sessions and workshops where city officials present initiatives and provide space for residents to voice concerns, have questions answered, share ideas, or provide input is vital. Fostering a collaborative and interactive environment where residents can contribute to tangible changes helps to build and maintain trust.

Cities can also give residents a formal platform for input through community advisory panels or committees. These committees can review how data is used, give feedback based on community engagements, and suggest improvements.

Finally, participatory budgeting opportunities that allow residents to vote on which initiatives get implemented and on what funds should be spent or allocated based on insights from shared data, drives citizen engagement. This approach demonstrates to residents that their contributions not only influence policy but also spending priorities. For example, participatory budgeting was [implemented in Paris](#), which was and remains the largest implementation in the world, and allows citizens to vote on projects designed and submitted by other citizens. According to an [interview with a former Cabinet Director Julien Antelin](#), participatory budgeting built both trust and public excitement.

By implementing these transparent, two-way relationships with residents, cities help residents see the positive, data-driven changes in their communities, and increase their willingness to participate in future initiatives.

## Additional Challenges to Successful Implementation

A major challenge to overcome in implementing such data-driven frameworks is managing the digital divide and ensuring that datasets are complete, meaning cities have sufficient data that is reflective of all groups and subgroups within their population. Residents from low-income or underserved communities may not have access to the range of digital devices that wealthier residents do, leading to uneven representation of data across the city and policies that don't address their needs. To overcome this, cities can take several approaches. First, by providing subsidies or loan programs for digital devices in underrepresented communities. Second, cities could collect data manually from engagement sessions and workshops held in those underserved areas. Third, cities could establish public data collection points, such as implementing air quality or noise monitors in those neighborhoods. New York City's [LinkNYC](#) project offers free Wi-Fi kiosks throughout the city and also allows residents to access public services announcements, among a range of other available features.

A second challenge to overcome is that of data overload and analysis capacity. An influx of data from thousands or even millions of devices could very quickly overwhelm city infrastructures, making it challenging to extract actionable insights in a timely manner. Advancements in artificial intelligence (AI) and machine learning (ML) could make it possible for cities to automate data analysis and highlight trends and anomalies. Additionally, establishing partnerships with local universities and other research institutions as well as tech and other private sector companies can expand analytical capacity. [Smart Nation Singapore](#) is a clear example of how cities can utilize AI-driven data analysis to manage and interpret large-scale data in real-time, helping city planners make quicker, data-informed decisions.

## Conclusion

In conclusion, establishing a voluntary, privacy-conscious, and citizen-driven data sharing initiative, offers an innovative and transformative approach to urban governance and community empowerment, while enabling cities to respond to residents' needs through better informed policies. Through this model, cities affirm the importance of individual privacy and trust, while creating robust data ecosystems that draw from a range of resources and expertise that are diverse and ensure that data insights are comprehensive. Further, it enhances inclusivity, making it possible for communities, regardless of technological access or literacy to contribute to and benefit from data-informed policies.

By addressing these issues, cities can build trust with residents and harness their data for public good. Success stories from Barcelona, New York City, Amsterdam, Singapore, and Paris show the potential impact of these initiatives, from improved air quality

monitoring to more inclusive urban planning and economic and social policies. As digital devices become more ubiquitous, embracing collaborative data-sharing frameworks could redefine how cities create policies, paving the way for healthier and more responsive urban environments.

Ultimately, as cities adopt this or similar frameworks, they have the opportunity to create communities where urban policies are more inclusive and equitable. By leveraging citizen-driven data and fostering collaboration, cities can meet today's challenges with solutions that respect resident autonomy and strengthen community resilience. This framework could thus serve as a blueprint for cities worldwide, to enhance urban life in ways that are both impactful and sustainable.

## References

Amsterdam Economic Board, *Amsterdam Smart City*. <https://amsterdamsmartcity.com/>

Chicago Data Portal, City of Chicago, *Array of Things Locations – Map*.  
<https://data.cityofchicago.org/Environment-Sustainable-Development/Array-of-Things-Locations-Map/2dng-xkng>

*Décider Pour Paris*, <https://decider.paris.fr/decider/jsp/site/Portal.jsp>

Institut Municipal d'Informàtica, *City Os*.  
<https://ajuntament.barcelona.cat/imi/en/projects/city-os>

Louisy, Khahlil and Zlatev, Juilan. *Insights: Digital Vaccination Credentials and The Public Trust - A U.S. Survey*, Institute for Technology & Global Health. 2021.  
[https://drive.google.com/file/d/1RTQWLzx5W6RMQEC75oYc4uVwJhnDoKCw/view?hsmi=129904957&hsenc=p2ANqtz-8lw9OU0fL-a6u\\_j7YR-vFDy5dlqdxL6kpeqpcWye6HPRIIsIHLgPYuimBK3QeSUDlK8iOfxK4GttfnE7pk0KZJVka\\_2Ciek\\_r2sNrG15ZKukIcD2Neg](https://drive.google.com/file/d/1RTQWLzx5W6RMQEC75oYc4uVwJhnDoKCw/view?hsmi=129904957&hsenc=p2ANqtz-8lw9OU0fL-a6u_j7YR-vFDy5dlqdxL6kpeqpcWye6HPRIIsIHLgPYuimBK3QeSUDlK8iOfxK4GttfnE7pk0KZJVka_2Ciek_r2sNrG15ZKukIcD2Neg)

*MyData*, <https://mydata.org/>

NYC Office of Technology and Innovation, *Artificial Intelligence*.  
<https://www.nyc.gov/content/oti/pages/reports>

NYC Office of Technology and Innovation, *LinkNYC*. <https://www.link.nyc/home.html>

Participatory Budgeting Project, *Expanding the imagination of democracy: PB in Paris*. June 13, 2017. <https://www.participatorybudgeting.org/pbparis/>

Rusama, Mikko. *Lessons from the City of Helsinki: Three Paradigm Shifts in Smart Cities*. MyData, Jun 14, 2024. <https://mydata.org/2024/06/14/lessons-from-the-city-of-helsinki-three-paradigm-shifts-in-smart-cities>

*Sentilo BCN*, <http://connecta.bcn.cat/connecta-catalog-web/>

Smart Nation and Digital Government Office, *Smart Nation Singapore*. Last updated November 20, 2024. <https://www.smartnation.gov.sg/>



 DATA-SMART  
CITY SOLUTIONS

**A PUBLICATION OF**

**Data-Smart City Solutions**

79 John F. Kennedy Street Cambridge, MA 02138

**[datasmart.hks.harvard.edu](http://datasmart.hks.harvard.edu)**

---



**BLOOMBERG  
CENTER FOR CITIES  
HARVARD UNIVERSITY**