41st WEDC International Conference, Egerton University, Nakuru, Kenya, 2018

TRANSFORMATION TOWARDS SUSTAINABLE
AND RESILIENT WASH SERVICES

Initial lessons learned in making sanitation data useful for decision making

JP. Archer and L. Noakes (United Kingdom)

PAPER [OFFICE USE ONLY]

Useful data is important for effective decision making. In 2017, the World Bank reported that poor data has led to an ineffective allocation of resources across the sanitation sector. The current data gap is also preventing the sanitation sector from tracking its progress in providing sanitation for people living in cities. This challenge is not insurmountable. A sector-wide response needs to include behaviour change in the collection, sharing and use of data. This paper explains Gather's approach to improving the usefulness of data through geospatial analysis, and presents our recommendation for a sector-wide data strategy that starts with the creation of a data standard for urban sanitation data.

Introduction

Gather was launched in 2016 as a research project to understand what was preventing the expansion of sanitation services in low income communities. In our first six months, we conducted interviews with nearly 100 professionals across a broad range of organisations working to make urban sanitation more sustainable and resilient across east Africa and south east Asia. Our research revealed that a lack of quality data on urban sanitation had made the sector less effective, efficient and collaborative than it could be. We also noticed an acute need for the sector to better track its progress in providing sanitation for people living in cities. We discovered that organisations providing sanitation services to the urban poor were collecting large quantities of data on their own operations, but that there was no mechanism to share and analyse this data to better understand the current state of sanitation across cities.

In 2017, we wanted to explore how we could improve access to urban sanitation data and geospatial analysis. We tested two approaches to learn about data collection and data sharing. First, we mapped shared sanitation infrastructure in Mathare in Nairobi, Kenya. We wanted to understand the costs and challenges of collecting, curating and analysing sanitation data. Second, we focused on the creation of a data sharing platform. We wanted to understand what was needed for sanitation providers to share the data they were already collecting, so that it could be analysed and the insights shared with sanitation providers and funders. Both approaches revealed the need for a sector-wide data standard for urban sanitation data. Without a data standard, it is impossible to collect and share data so that it can be analysed to identify where investment is needed to expand sanitation services.

Methodology

Our methodology is built on a collaborative approach to systems change. As well as continuing to listen to the experiences and needs of the sanitation providers we had interviewed, we also want to engage with experienced partners in the fields of data collection and data sharing.

In August 2017, we partnered with Spatial Collective – a local, respected social enterprise – to collect data on shared sanitation in the informal settlement of Mathare in Nairobi, Kenya. Mathare is home to an estimated 180,000 people. A team of eight young local enumerators spent two weeks collecting data on the GPS coordinates and 17 additional attributes for every communal toilet in Mathare. The team captured data on a total of 119 communal toilets. We then analysed this data set alongside openly accessible data on population density from WorldPop at the University of Southampton.

In June 2017, we began to design the architecture for an urban sanitation data platform with the support of Geovation, the UK Ordnance Survey's location data hub. A platform enables value-creating interactions between two groups of users.

Approach 1: Improving availability of data by mapping sanitation infrastructure

We faced two challenges when considering how to execute and expand this approach. The first was the limited published guidelines on what indicators should be collected and how. Additionally, we faced poor economies of scale if we wanted to expand our survey to cover the entirety of Nairobi or to replicate the model in other cities inside or outside of Kenya. Our interviews with sanitation providers showed us that these challenges were shared across the sector.

Traditional data collection methods, such as household surveys, can be helpful for snapshot assessments. For example, our analysis revealed that only 3% of the population of Mathare had access to basic shared sanitation. However, we quickly discovered the limits of this approach for strategic decision making: the data quickly goes out of date, and regularly updating it is prohibitively expensive.

Approach 2: Improving access to data analysis by building a data sharing platform

We were able to perform some basic analysis on shared sanitation in Mathare from the data that we had collected. To understand the state of sanitation across the whole of Nairobi though, we needed to supplement our data collection with data from sanitation providers working in the city. These organisations also wanted to access city-wide data sets, but there was no system available that allowed them to do this. We concluded that we needed to focus on improving the ability of sanitation providers to share their data. This led to our second approach: a data sharing platform.

In the early designs of our platform, we envisioned that organisations providing sanitation would upload data on their services. Gather would then perform geospatial analysis on this data and publish it as an interactive map. Finally, sanitation organisations would then use the map to make decisions on which parts of the sanitation value chain in a city they should prioritise for investment.

We started by launching two demos of our platform, showcasing the data we collected in Mathare in partnership with Spatial Collective. The first demo visualised the sanitation data and allowed the user to interact with it, e.g. display toilets connected to a sewer with handwashing facilities. The second demo visualised analysis we had performed on the sanitation data alongside other data sets including population density, e.g. show where new toilets should be built, or where hand washing facilities should be installed. The purpose of these demos was to demonstrate our concept to sanitation providers so that we could work with them to design a fully operational platform where they could easily share and learn from data.

One early success of the launch of these demos was the interest we received from organisations providing sanitation. They offered us data sets on sanitation services in five cities to contribute to the building of our full working platform. We had previously assumed that these organisations would only share their data with us if major funders of sanitation were also signed up to the platform. We had underestimated the strength of another incentive: these organisations wanted to understand what their data could tell them, but often did not know what questions to ask their data or how to ask them.

When we designed the demos, we had assumed the technology would be a significant challenge. Our workshops with Geovation (and a separate project designing a crisis response tool with IBM Research) taught us that readily available geospatial software and programme architecture could perform the analysis we required. However, when we received the first data sets from sanitation providers, we discovered that in fact our biggest challenge was the quality of the data that we had been sent. There was little uniformity across the data sets, apart from the GPS coordinates of the toilets. Furthermore, the surveys had frequently been designed with limited consideration of the analysis that would need to be performed on the data. We could perform exploratory analysis on the data sets (like the analysis on our data from Mathare) but we were unable to perform deeper analysis on the individual data sets or perform comparative analysis between them.

This showed us that creating and implementing standards for sanitation data would be fundamental to the success of our data sharing platform. The lack of standardisation across the data we received made it necessary for us to recommend a new data standard for the collection and management of urban sanitation data before we could proceed with building a data sharing platform. In March 2018, we hosted a data dive to explore this new data standard for urban sanitation. Twenty data scientists analysed the data that sanitation providers had shared with us. The data scientists investigated the validity and reliability of the data, how the data could be used to estimate the coverage of sanitation services and assess health risks to the population, and how the data could be used to optimise waste collection routes. The data scientists were able to create

geospatial visualisations of the data to provide insight, but the variable quality, format and structure of the data confirmed the need for a sector-wide strategy for data collection and management.

Lessons learned from different approaches to making sanitation data useful

Data standards are crucial. It is very difficult to properly compare or collate data sets when the data has been captured in different formats through different surveys that have relied on different questions and definitions. We can see from the data sets that were shared with us that the sanitation sector predominately collects data on operations. This focus can be useful in understanding an individual organisation's operations. However, it does not allow the sector to share data and understand the level of sanitation in a city. This makes providing sanitation in a strategic, collaborative way much more difficult.

The lack of data standards and ability to analyse data is not unique to the urban sanitation sector. In March 2017, the United Nations Statistical Commission's Cape Town Global Action Plan for Sustainable Development Data highlighted the necessity for data standards to enable comparative analysis of data, and noted the potential for geospatial analysis to provide greater insight. In November 2017, the Inter-Agency and Expert Group on SDG Indicators Working Group on Geospatial Information noted that data availability remained a key challenge and that members were currently unable to develop sufficient analytical capacity in-house. We believe these challenges are not insurmountable, but will require a collaborative effort by non-governmental practitioners to support national statistic offices and key decision makers.

The solution lies in standardising key indicators (including the format and definitions) that are collected by the urban sanitation sector. Creating a data standard for urban sanitation data would also provide an opportunity to align city indicators to the country indicators set by the Joint Monitoring Programme (JMP) for monitoring progress on Sustainable Development Goal 6.2. It is important to note one of the key challenges that will need to be addressed when establishing uniform definitions for indicators: the JMP definition does not currently include shared sanitation in its definition of safely managed sanitation. Shared sanitation is very common in informal settlements, particularly when space is at premium and there is limited room for household sanitation. Members of the JMP team have concluded that there is no reason to stop investing in shared sanitation, so any future data standard should include data on services that are defined as shared sanitation.

Conclusion

The urban sanitation sector needs to adopt a sector-wide data strategy, and this starts with the adoption of a data standard for urban sanitation. A data standard would ease the burden and cost of data collection on sanitation providers by simplifying the data they need to collect and share and reducing duplication of efforts. Standardised data, accessible through a data sharing platform, will allow organisations across the urban sanitation sector to work more effectively, efficiently and collaboratively to provide sanitation to people living in low income communities.

References
EVANS, B ET AL 2017 *Editorial: Limited services? The role of shared sanitation in the 2030 Agenda for Sustainable Development* Journal of Water, Sanitation and Hygiene for Development 2017, pp 1-3
PARKER, G.G., VAN ALSTYNE, M.M and CHOUDARY, S.P 2016 *Platform Revolution: How Networked Markets are Transforming the Economy and How to Make Them Work For You*. W. W. Norton & Company, New York.
UNITED NATIONS STATISTICAL COMMISSION 2017 *Cape Town Global Action Plan for Sustainable Development*. Prepared by the High-level Group for Partnership, Coordination and Capacity-Building for Statistics for the 2030 Agenda for Sustainable Development.

WORLD BANK GROUP 2017 *Reducing Inequalities in Water Supply, Sanitation, and Hygiene in the Era of the Sustainable Development Goals: Synthesis Report of the WASH Poverty Diagnostic Initiative*. WASH Synthesis Report. World Bank: Washington, DC.

Contact details

John Peter Archer and Lindsey Noakes founded Gather in 2016. They were included in Forbes 30 under 30 social entrepreneurs in Europe 2018. John Peter and Lindsey share a passion for the application of data analysis to strengthen and transform systems that serve the most marginalised in urban communities.

John Peter Archer	Lindsey Noakes
Gather	Gather
Urban Innovation Centre, 1 Sekforde Street, London, EC1R 0BE	Urban Innovation Centre, 1 Sekforde Street, London, EC1R 0BE
079851966846	07986086444
john@gatherhub.org	lindsey@gatherhub.org
www.gatherhub.org	www.gatherhub.org