

How City Data Hubs with Data Standards will create actionable insight to accelerate and track progress on SDG 6.2



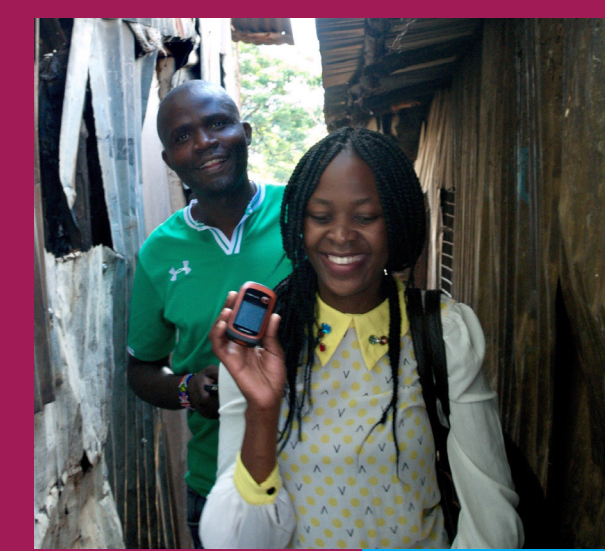
1. The Data Gap

The best solutions require information, collaboration and targeted investment. This is impossible without data.

The urban sanitation sector suffers from a major lack of useful data that it can use to drive decision making. In 2017 the World Bank reported that poor data had led to an ineffective allocation of resources across the sanitation sector. In 2018 the United Nations found that the current lack of quality data had limited the sector's ability to monitor its progress, and stated that "reliable, consistent and, whenever possible, disaggregated data are essential to stimulate political commitment, inform policymaking and decision-making."

We spent twelve months interviewing over one hundred organisations. The majority were trying to bridge the data gap by conducting independent surveys. This allowed them to create snapshot baselines and track progress specific to their service. These efforts were repeated across the sector, but there was no mechanism to share this data to create a regularly updated resource that could inform collaboration and investment into infrastructure and services across an entire city.

We decided to test three different approaches to bridging the data gap.



2. Mapping Sanitation

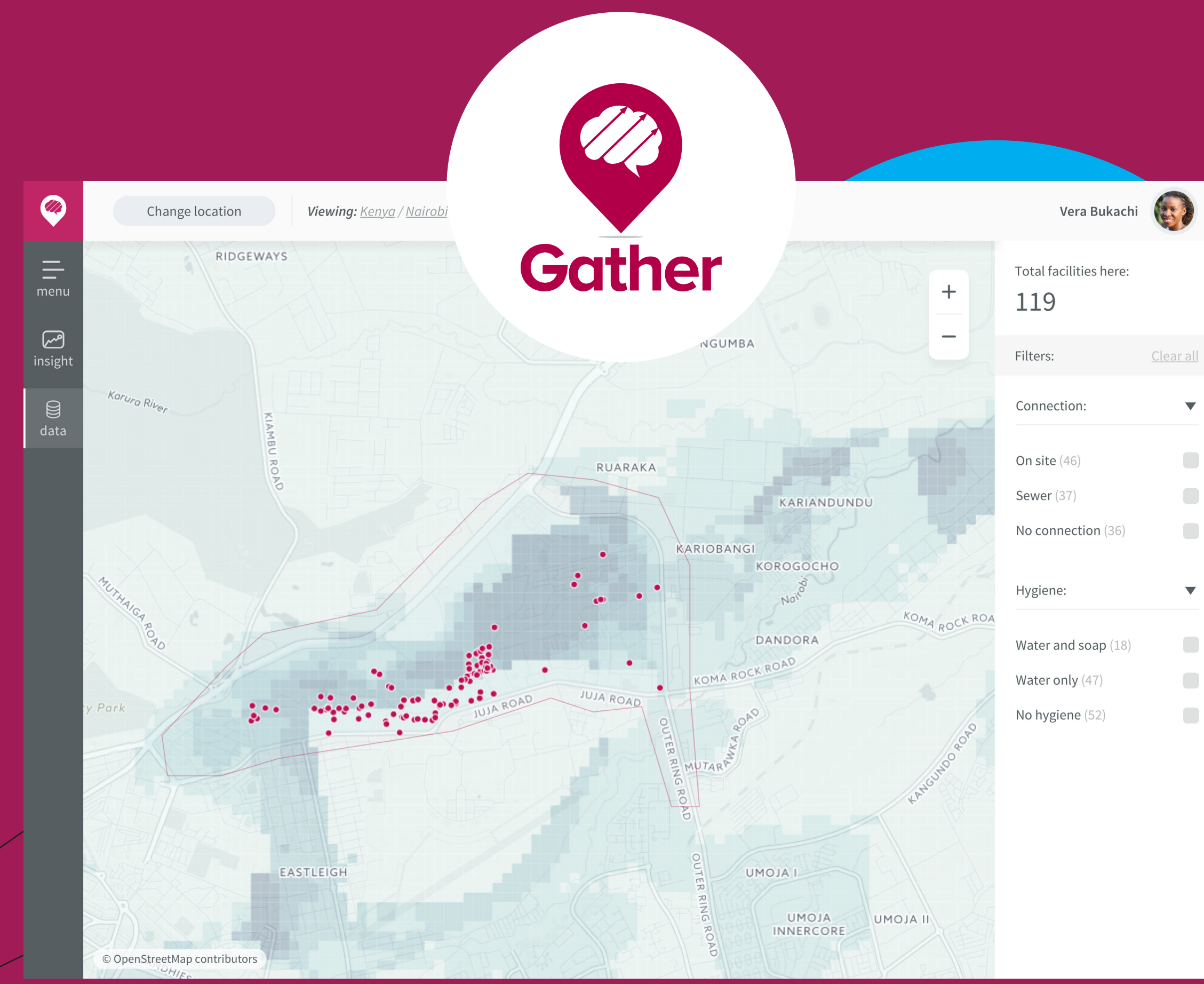
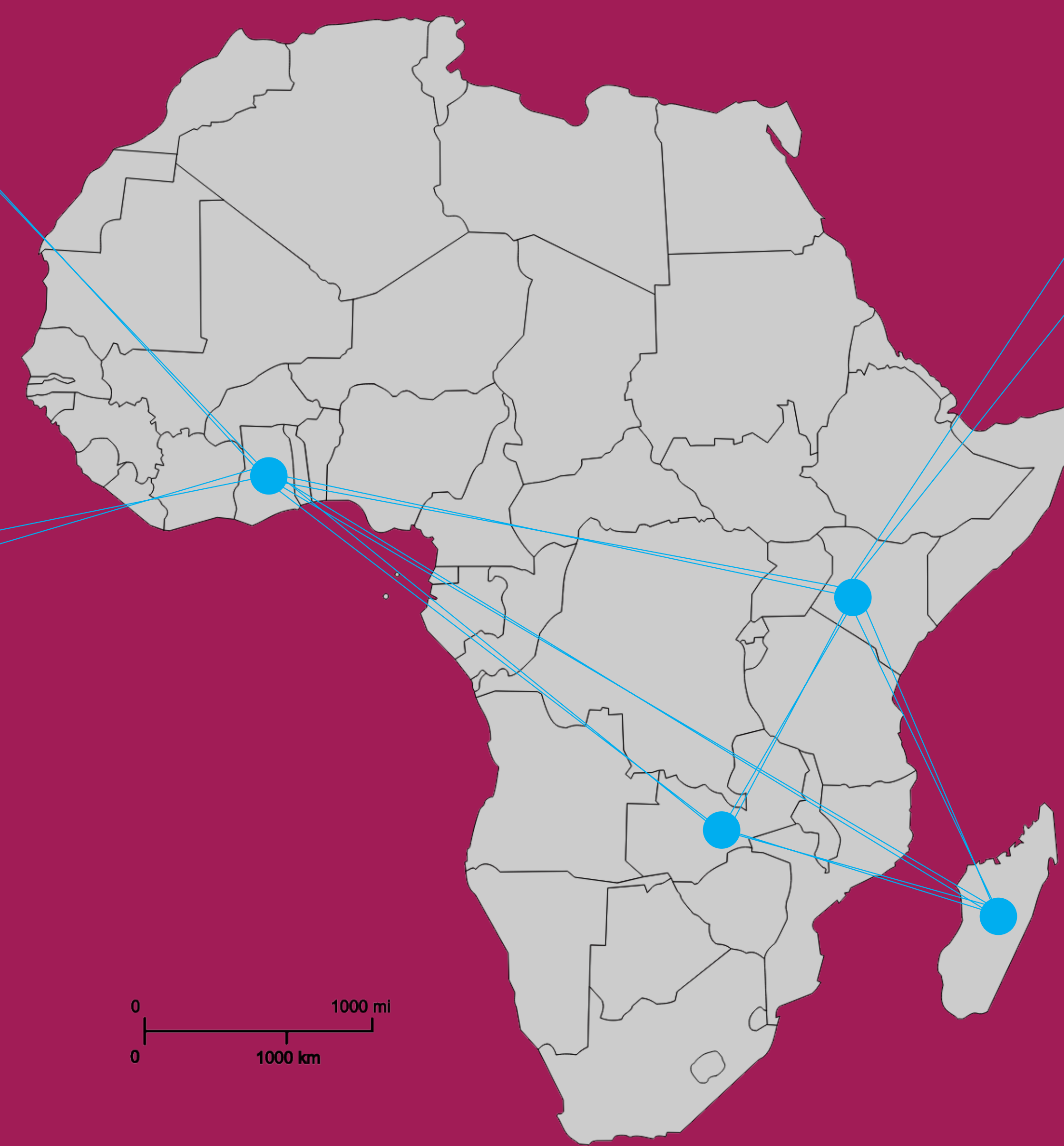
Successes: We partnered with Kenyan social enterprise Spatial Collective to map communal and household sanitation in Mathare, Nairobi. A team of eight young local enumerators collected data on the GPS coordinates and 17 attributes for the 119 communal toilets in Mathare. We then analysed this data set alongside open data on population density from WorldPop at the University of Southampton. Our analysis revealed that only 3% of the population of Mathare had access to basic shared sanitation.

Challenges: We faced poor economies of scale if we wanted to expand our survey to cover the entirety of Nairobi or replicate the model in other cities. Traditional data collection methods can be helpful for snapshot assessments but the data quickly went out of date.

3. Building a data sharing platform

Successes: We partnered with the Ordnance Survey's Geovation Hub to design an online data sharing platform. We designed it so that organisations could upload data on their services and then access interactive, geospatial analysis on the state of the sanitation value chain in their city. We also explored options for language-based queries and mobile data with IBM Research.

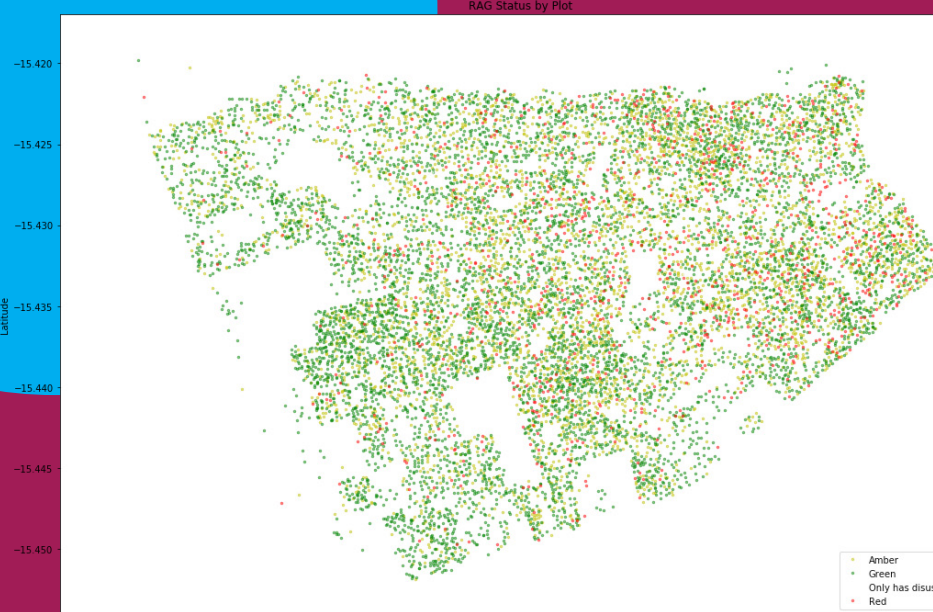
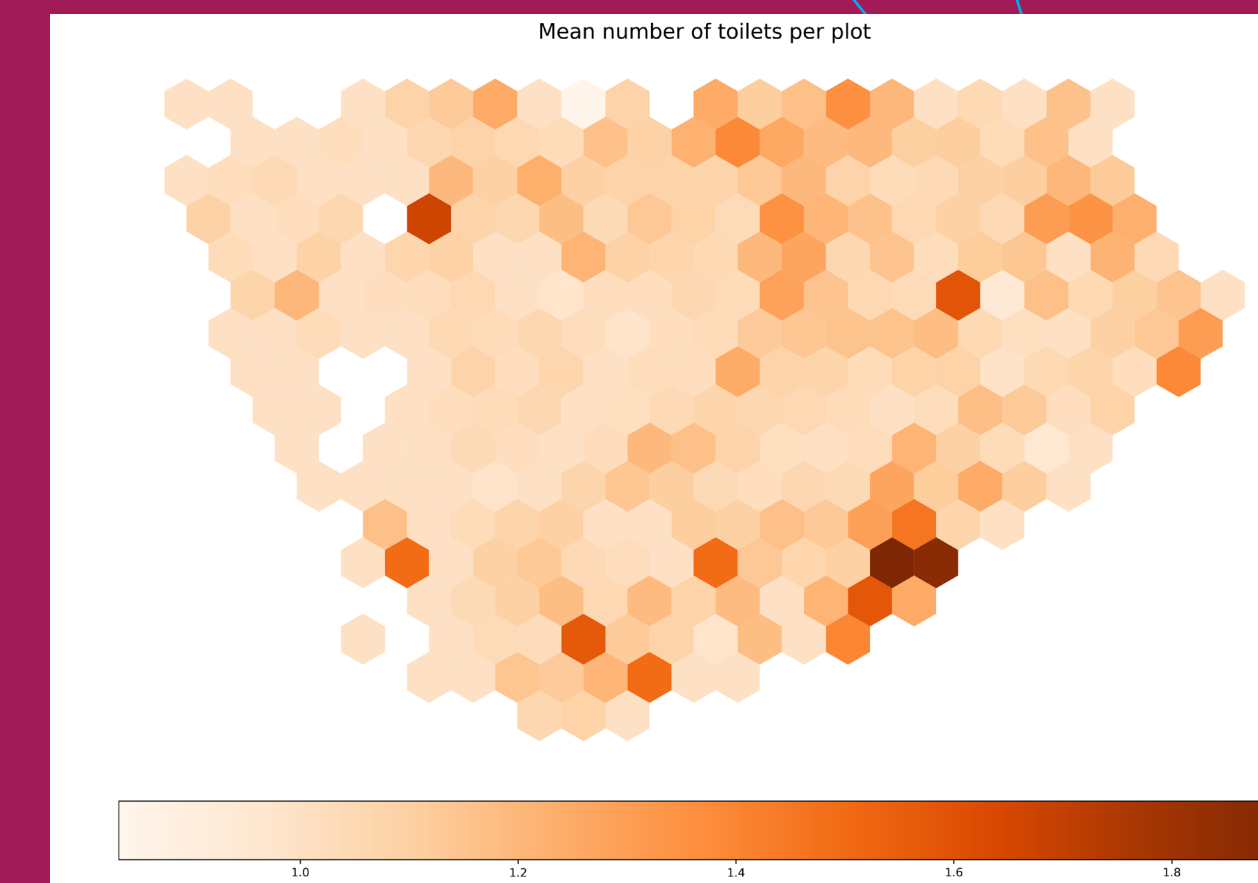
Challenges: We had assumed that technology would be a significant challenge. However, we discovered that the biggest challenge was the quality of the data. There was little uniformity in how data been collected and managed. There was even disparity in how location of infrastructure had been recorded. We were able to perform exploratory analysis but unable to perform more useful, prescriptive analysis.



4. Data Analysis

Successes: 20 data scientists and academics joined us for a two-day data dive to explore toilet data from Lusaka, Zambia. The team assessed the quality of the data and created a series of maps to visualise information on the mean number of toilets per plot, disused toilets and risk status per plot. A key piece of work was the creation of zoning vectors to visualise the information. The team also started to perform prescriptive analysis to see if the data set could be used to optimise waste collection routes and advise on priority areas for investment.

Challenges: The data set was one of the most comprehensive we have seen. We had hoped to perform comparative analysis with data sets from other organisations but there was still a significant disparity between how the data have been collected, managed and cleaned.



6. City Data Hubs

We plan to launch a city data hub in Antananarivo, Madagascar in early 2019. We want to prove the importance and value of data practices, data standards and data driven decision making. We also want to show that local organisations can collaborate to create a solution to the sanitation data gap in Antananarivo and take action to end the sanitation crisis.

Why Antananarivo? One of the aims of the Madagascar government's Integrated Sanitation and Drainage Master Plan for Antananarivo is to better define and prioritise investment needs for the sanitation sector. This requires geospatial data of higher resolution than currently available. Also, the Madagascar Ministry of WASH's existing database Suivi Eau et Assainissement à Madagascar hosts data on sanitation infrastructure, but still lacks data from urban areas.

We have designed the city data hub to overcome several key challenges expressed by each of the hub members. The first is the data gap: they cannot access the trusted, actionable data they need to understand the areas they provide services, plan strategically or identify new markets. The second is the need for a physical space to learn best practice on sanitation data, discuss their data challenges, and collaborate with their colleagues in the sector. Finally, there is the data burden: these organisations commit significant time, funding and energy to collecting and analysing data in ways that often feel inefficient or repetitive.

One of the key outputs of the project is the geospatial shit flow diagram. We will publish a visualisation of this to make it useful to the wider sector working in Antananarivo, but to gain full access to the data, organisations must become a member of the city data hub. In this way, we can ensure the SFD remains updated and the hub ecosystem continues to grow.

5. Data Standards

Data Standards are not glamorous, but they are essential. They are often developed by subject matter experts to reach a consensus on how to solve problems and allow for common practice in the collection, management and exchange of data.

Data standards enable transparency and understanding. They provide consistency for well understood terms, codes and data structures including common attributes and definitions. This allows for comparisons or shared understanding even when all the data is not standardised.

We can see from the data sets that have been 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.

Gather is looking to launch a Data Group of experts from across the sanitation sector and data sector to help create and maintain a new set of data standards for urban sanitation. If you are interested, please contact us.



@gatherhub_
hello@gatherhub.org
gatherhub.org

Awards 2018



Partners GEOVATION



Deloitte. generation