**38th WEDC International Conference, Loughborough University, UK, 2015**

WATER, SANITATION AND HYGIENE SERVICES BEYOND 2015:
IMPROVING ACCESS AND SUSTAINABILITY

# Tapping customers: a spatially-explicit, open-source platform for crowdsourcing water service data in Ghana

*J. Stoler, C. Ewing, R. Appiah Otoo, L. Tran, N. T. Barimah Adu,
M. Nyoagbe & D. N. Okai Nunoo (USA & Ghana)*

**BRIEFING PAPER 2090**

*Municipal water rationing has become increasingly common in developing urban centres, leading to substantial variation in service levels among residential customers. This paper introduces an open-source tool called Improving Quality of Urban Water Service by Engaging SMS Technology (IQUEST), a geographic decision support system that harnesses crowdsourced water data to enable Ghana Water Company Limited (GWCL) to monitor residential water service quality in Accra, Ghana. This paper presents the conceptual model, general architecture, and user interface, and highlights the potential for other municipal water managers to implement a decision support tool in a resource-constrained setting.*

## Introduction

In 2014 an estimated 748 million people lacked access to an improved drinking water source (WHO/UNICEF, 2014). The burden of water- and sanitation-borne diseases is increasingly an urban challenge that is expected to intensify over the next decade. Sub-Saharan Africa is unlikely to meet its Millennium Development Goals for drinking water provision, and contains almost all of the nations in which a minority of the population had access to an improved drinking water source as of 2008. It is the only region where urban access to piped water (either in the home or at a public tap) *decreased* over the last two decades, from 68% in 1990 to 55% in 2008 (WHO/UNICEF, 2011).

Accra, Ghana, epitomizes many of the challenges faced by developing cities in the effort to provide basic services; the Accra Metropolitan Assembly simply has not been able to maintain—much less expand—public infrastructure fast enough to keep up with its growing urban population. This imbalance of supply and demand is typical of low-income nations and fed by both the legacy urban structure imposed by colonialism, and by continued in-migration of poorer rural migrants (Nickson & Franceys, 2003). Water shortages have traditionally been attributable to both production limitations and poor water resource management (Nsiah-Gyabaah, 2001), and consequently Ghana Water Company Limited (GWCL) implements a rationing program that manages limited water resources. This rationing program is inadvertently influenced by locally-varying geography and income levels, and results in substantial neighbourhood-level heterogeneity of water delivery (Stoler, Weeks, & Appiah Otoo, 2013).

Data on water service levels is traditionally provided by GWCL water districts based on the local rationing schema, but this 'top-down' data is considered an inadequate assessment of household water service. Many factors adversely affect water service, such as pressure, elevation, colour, and odour. Although daily water quality checks are performed at selected nodes on the main distribution network by the central GWCL laboratory, none of these factors are systematically recorded from water within the service lines. In 2010 GWCL and Kwame Nkrumah University of Science and Technology piloted a short door-to-door survey of 1,876 Central Accra households to measure water service levels in a 'bottom-up' fashion using a metric developed with World Bank consultants called the distribution quality factor ($fQ$) and managed in a geographic information system (GIS) (unpublished data). The $fQ$ metric is an aggregation of responses to water quality survey responses at the neighbourhood level. The pilot demonstrated variability in neighbourhood water service at a finer spatial scale than is reported by the water districts, and these results

elucidated the potential for spatial modelling of service levels elsewhere in the city. While the *f*Q methodology was promising, it was not pursued further because the household survey was too expensive and time-intensive for full implementation, and the pilot study area was inappropriate for broader statistical extrapolation due to its mixed residential and commercial profile. GWCL thus began to consider how the integration of new communications technologies could mitigate the cost of customer service data collection.

Short message service (SMS) is a common and inexpensive communication vehicle which has long held potential for improving health outcomes in the developing world (Kaplan, 2006). A number of reviews highlight the growing use of mobile health (mHealth) applications such as SMS in the provision of health services (Free et al., 2013; Gurman, Rubin, & Roess, 2012; Krishna, Boren, & Balas, 2009). Despite this progress, a number of significant challenges remain to implementing mHealth applications in resource-constrained settings, and these range from variable literacy and socio-cultural patterns among the target population to telecommunications service provider instability (Arul, Holley, & Bas, 2013; Medhi et al., 2011; Siedner et al., 2012). These challenges have slowed the widespread adoption of SMS and automated dialling systems to little more than another marketing or customer service tool for public utilities or government services in developing economies.

This paper describes a new tool called *Improving Quality of Urban Water Service by Engaging SMS Technology* (IQUEST), a geographic decision support system that aims to harness crowd sourced water data for monitoring residential water service quality by GWCL at a significantly reduced cost. The project uses an SMS gateway to solicit, collect and parse data, and an open-source interactive data mapping software to drive an analytic management dashboard. The integration of data from multiple platforms exemplifies an increasingly data-driven approach to water management (Kapoor & Menon, 2014), while recognizing mobile technology's role as a new cornerstone of sustainable public participation (Ochara & Mawela, 2013). This pilot, scheduled for late 2015 and limited to the Accra Metropolitan Area, will inform the potential for engaging Accra's public and ultimately rolling out IQUEST to other areas in Greater Accra Region served by GWCL. IQUEST has the potential to inform district managers of possible leaks and/or water piracy, identify areas that are in particular need of non-emergency infrastructure maintenance, and ultimately inform water resource management. The IQUEST code is open source and may be portable to other developing-world urban centres with residential water monitoring issues.

## System components

The participation of water customers in reporting water service quality over time is central to the data collection process for IQUEST. Pilot interfaces are planned for three types of incoming data streams: solicited SMS, volunteered SMS, and solicited voice. Outbound text messages are generated by IQUEST, and all messages (inbound and outbound) are transmitted by the SMS gateway called TXTMeGhana. This service provides an application programming interface (API) leased by telecommunications companies as a dedicated channel to accept SMS messages, in our case on behalf of MTN, the mobile provider for GWCL. TXTMeGhana passes incoming response data to the parsing module for conversion and storage within IQUEST. Voice data are solicited and collected by an interactive voice response system (IVRS). IQUEST then aggregates all submissions and displays them in charts available to GWCL management via an online dashboard, which is the focus of this paper. The general system architecture for IQUEST and related interfaces is summarized in Figure 1.

The initial pilot of IQUEST will pursue four indicators of customer water quality: colour, particles, taste/odour, and hours of flow, and each data stream will crowdsource these indicators in a slightly different manner. The four indicators were selected from a longer list of indicators that were measured in the 2010 GWCL study that produced the *f*Q metric, and were found to have the strongest associations with the *f*Q. These methods of customer data collection present a nominal cost to the customer in the form of a minute of their time to send the SMS or answer a call, as SMS messages to GWCL are toll-free, and Ghanaian mobile plans do not charge for incoming calls. That said, all of these crowdsourcing strategies will require a local branding and marketing strategy with a community presence in order to repeatedly remind customers of the importance of the initiative, and the potential service improvements associated with their participation. These marketing efforts will be specifically targeted toward lower-income neighbourhoods, as these customers are more likely to experience water stress due to intermittent service. Higher-income families are more likely to own pumps and large water tanks in order to store water during periods of intermittent delivery. These families are generally less affected by rationing schedules and maintenance-related shutoffs, and therefore less likely to provide useful information about hours of flow.

**Figure 1. Architecture diagram for IQUEST showing data streams and database interfaces**
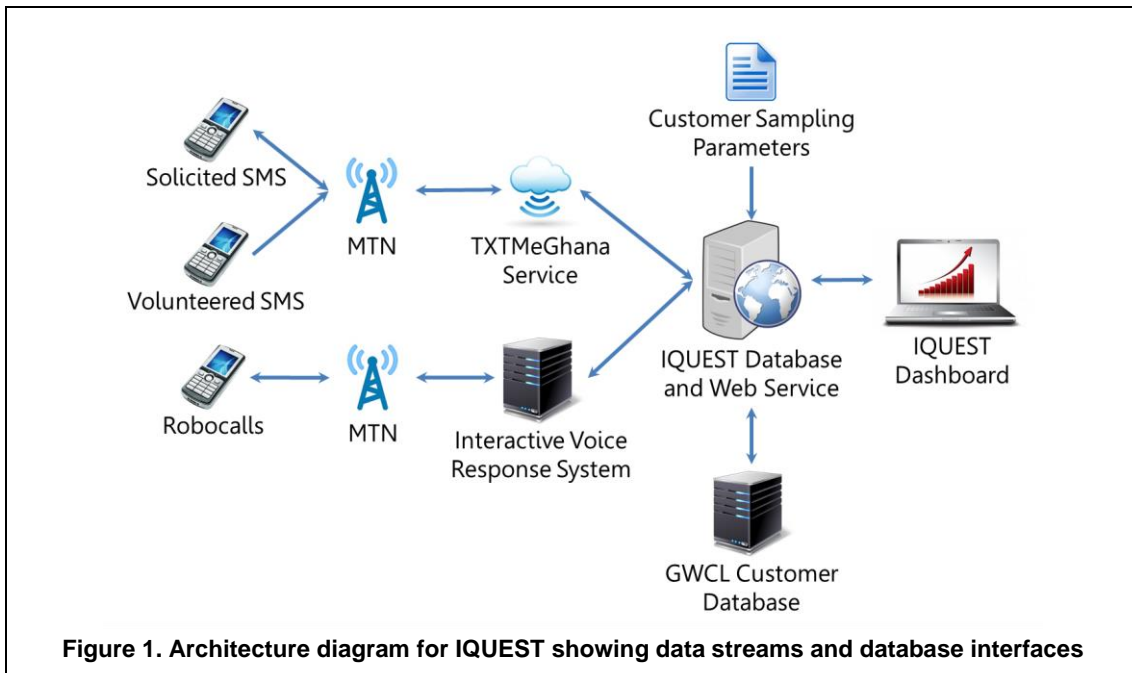
## Solicited SMS

The primary mechanism for crowdsourcing water quality data is the solicitation of SMS messages via SMS itself. A custom PHP interface allows IQUEST to perform three key functions using the GWCL customer database: search, send, and update. That is, the system will *search* customer records and apply a geodemographic sampling scheme to select target customers, *send* outgoing SMS messages to a selected group, and then *update* the customer record with the timestamp of when the customer was contacted. Customer queries will typically be based on geographic location (using neighbourhood codes), language spoken, and the date of last SMS inquiry in order to generate targeted messages that increase the probability of customer response.

## Solicited voice

Customers will also be sampled to receive brief automated telephone solicitations, i.e. robocalls, to volunteer information about the quality of their water service. The IVRS will be able to access the GWCL customer database to target geodemographic characteristics and ensure that the correct language script is presented to the customer. The customer would hear a short introductory script and be asked to press 1 on their keypad to opt into the survey. The IVRS then asks sequential questions about each indicator and specifies the numeric keys, i.e. touch-tones, which correspond to each valid response. The interactive nature of the IVRS allows all four indicators to be collected in a single call that would last no more than a minute or two. After a customer completes a survey, the IVRS will conclude with a thank-you message that is similar to the one scripted to SMS participants, and they will not be contacted for at least another 30 days.

## Volunteered SMS

One of the shortcomings of targeted solicitations via SMS or automated calls to GWCL customers is that the customer database typically only contains contact information for one member of a given household. The individual on record may spend much of their day away from the home, and may not be the most reliable source of information about water service. GWCL therefore wants to democratize the reporting process by providing a mechanism for other household members—or any community member in a GWCL service area—to proactively send us information about water service issues. Project managers have designed colourful postcards that instruct water customers to submit short, coded water reports via SMS (toll-free) that capture the basic water flow and quality information described above.

To send a water report, a customer will enter a numeric neighbourhood code plus descriptors of yesterday's water service to the water company's dedicated SMS short code. The message then goes into a central database where it is parsed into a neighbourhood code and four water quality indicators and a date timestamp. The IQUEST system then aggregates all submissions for display in charts and maps via the

management dashboard. To motivate customer submissions, the instructions will explicitly express GWCL's interest in hearing from its customers in order to improve water service. As an incentive during the pre-launch, GWCL plans to advertise a monthly drawing in which the company randomly selects a participant from each pilot community to receive free phone credits.

### Input parsing

The parsing module processes customers' text messages into a format suitable for queries from the IQUEST dashboard application. For solicited SMS, the parsing module first separates the customer's mobile number from the message header and matches it to the customer's neighbourhood by querying it from the GWCL customer database. For volunteered SMS, the parsing module interprets the numeric neighbourhood code from the first few characters of the message body. In either case, the parsing module then separates the message body into characters and keywords and attempts to match them to a particular indicator category (colour, particles, taste/odour, and hours of flow) using a validation algorithm. If a response is valid, the timestamp, mobile number, and indicator value is logged to a new record in the IQUEST database. The customer-sourced data is now available for querying.

## Prototype

The core of IQUEST is a web-based application written in PHP and JavaScript with a MySQL database. Incoming message data is processed by IQUEST's web service and stored in a MySQL database where it is accessed through the online dashboard. This section describes the web service and dashboard, and summarizes current and planned features of IQUEST's user interface (UI).

### Web service and dashboard specifications

A simple API written in PHP handles requests between IQUEST and each system interface; the web service is essentially the translator that allows each system component to communicate with each other. The web service provides the engine for querying the IQUEST database using parameters such as a date range or neighbourhood location. It also returns data that are properly formatted for display in the dashboard using Google Charts. The dashboard utilizes various open source projects to provide a simple and effective tool for management. The presentation layer is handled by Twitter Bootstrap, which allows the website to easily respond to different devices and screen sizes. In addition to the dropdown menu of neighbourhoods, a map allows users to navigate neighbourhoods spatially by loading KML through OpenLayers. The water quality data is retrieved through AJAX calls using jQuery and presented through Google Charts.

The PHP web service script and dashboard source code are open source and available at the following Internet address, along with screenshots of the UI: https://github.com/claytical/IQUEST/archive/master.zip.

### The management dashboard (user interface)

The dashboard is the user interface of IQUEST that is accessed by water managers at GWCL. The dashboard is designed for simplicity and ease of use, and the prototype consists of two primary views: a summary view with overview information for all of the neighbourhoods that are being managed by IQUEST, and a neighbourhood view containing a dynamic map, neighbourhood-level summary statistics, and charts of each indicator.

After username/password authentication, the user sees a summary view screen with two tabs, *Overview* and *Index Ratings*, which each contain summary data for all enrolled neighbourhoods. The *Overview* tab summarizes the number of customer responses received by IQUEST over a defined time interval in a line graph. The *Index Ratings* tab displays a bar chart of a composite index created by awarding points for positive-quality customer responses, and dividing the number of actual points by the total number of potential points accumulated if all responses indicated positive quality, thus scaling the value from 0 (worst) to 100 (best). On either tab, the user may change the reporting time interval to any one of five presets (e.g. past 30 days, past 3 months, etc.) or define a custom time interval. This view also contains a drop-down menu in the upper right-hand corner for selecting a specific neighbourhood and drilling down to a detailed data view of that area, i.e. the neighbourhood view.

When a neighbourhood is selected from the summary view, the user is taken to this neighbourhood's view (see Figure 3, presented with simulated data). This view presents a KML neighbourhood map of the Accra Metropolitan Area draped over Google Map tiles. The user can select a defined or custom time interval for data display (just as in the summary view), and review information specific to the selected neighbourhood.

The user is presented with counts of the total number of customer reports, the number of unique users reporting, the average and median number of reports per participating customer, and a snapshot of the most frequent value reported for each of the four water quality metrics. Below the neighbourhood summary, the user can toggle among four tabs displaying a stacked area plot of each water quality metric over the selected time interval.

Overall, the summary view gives the GWCL manager a broad snapshot of neighbourhood-level variation in both customer participation and the water quality index. This offers management a simple qualitative tool for identifying potential neighbourhood outliers in customer engagement and water quality, an expected improvement over the status quo of relying on anecdotal stories from water district managers or spurious customer service complaints. The neighbourhood view allows management to qualitatively explore geographic patterns among communities with water quality shortfalls, and compare individually crowdsourced water quality metrics by neighbourhood. This could highlight the source of water distribution problems related to pump failures, cracks and leaks due to infrastructure aging or piracy, and other issues that are often also sources of non-revenue water and opportunities for cost recovery by the utility. By visualizing the geographic proximity of subpar water quality reports, GWCL management will be able to schedule more targeted field visits and more effectively deploy maintenance and repair projects, given the utility's limited resources. The ability to more quickly identify network failures could even have a public health impact, as Accra has intermittently faced relapses of cholera associated with the co-location of shallow, compromised sewage and water infrastructure (Dzotsi et al., 2014).

### Future enhancements

IQUEST was developed with basic data views as a proof-of-concept, but there are a number of additional features that are queued for a future release. The primary architectural change is, broadly speaking, the ability to capture input from a more diverse set of social media inputs, particularly WhatsApp and Viber, as smartphone adoption increases in Ghana. For now, opportunities for improvement are linked to improving the user interface. GWCL also hopes to add a reporting module that allows a user to export any data selection into canned reports or comma separated value (.csv) format for use in other data analysis tools.

## Conclusion

Accra's urban water sector has experienced significant inertia since the 2011 transfer of operational authority from the previous operator, AVRL, to Ghana Urban Water Limited, a provisional structure that managed operations for two years before ultimately transferring power back to GWCL. While plans are underway to expand water production for the Accra Metropolitan Area, cost recovery by the utility remains a primary concern. New initiatives to implement pre-paid water meters have been met with public outcry, particularly over social justice issues related to serving vulnerable populations, and elevated negative public sentiment against GWCL as an extension of the government. In this environment, GWCL is motivated to improve operational efficiency with its existing human and technological capacities, and the IQUEST system is one small exploratory step toward a customer-oriented focus that productively engages customers.

IQUEST is an inexpensive, open-source system that requires basic database management and programming expertise, but is relatively simple to deploy. It can accommodate multiple streams of incoming customer data, both solicited and volunteered, and offers management a geographically-oriented view of customer feedback with the potential to create high-impact maps of custom water service quality indices. By engaging customers in a participatory approach, it has the potential to redeem the public image of GWCL in communities via the pairing of modest degrees of civic engagement with competent water infrastructure maintenance and repair. IQUEST may be portable to other developing nation contexts and potentially paired with other new mHealth applications that objectively measure water quality in the field (Kumar Gunda et al., 2014). It may be useful in other mobile-phone-heavy populations facing water management challenges, whether related to a municipal piped network, or to a centrally-managed constellation of boreholes. Accra's mobile-savvy population presents a generally untapped source of participatory data that may help provide clarity to an otherwise cloudy future of municipal water management in Ghana.

**References**

Arul, C., Holley, W. and Bas, H. (2013) *Vulnerabilities in mHealth implementation: a Ugandan HIV/AIDS SMS campaign*. Global Health Promotion Vol. 20, No Suppl 1, pp. 26-32.

Dzotsi, E., Asamoah, A., Noor, C., Gershon, A., Yirenchi, E., Nuoh, R. and Atelu, G. (2014) Report of cholera outbreak investigation, Greater Accra Region, June-August 2014 (pp. 26). Ghana School of Public Health: Accra.

Free, C., Phillips, G., Watson, L., Galli, L., Felix, L., Edwards, P., Patel, V. and Haines, A. (2013) *The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis*. PLoS Medicine Vol. 10, No 1, pp. e1001363.

Gurman, T.A., Rubin, S.E. and Roess, A.A. (2012) *Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature*. Journal of Health Communication Vol. 17, No Suppl 1, pp. 82-104.

Kaplan, W. (2006) *Can the ubiquitous power of mobile phones be used to improve health outcomes in developing countries?* Globalization and Health Vol. 2, No 1, pp. 9.

Kapoor, N. and Menon, G. (2014) *A new data paradigm for water management*. Water Utility Management International Vol. 9, No 3, pp. 16-19.

Krishna, S., Boren, S.A. and Balas, E.A. (2009) *Healthcare via cell phones: a systematic review*. Telemedicine Journal and e-Health Vol. 15, No 3, pp. 231-240.

Kumar Gunda, N.S., Naicker, S., Shinde, S., Kimbahune, S., Shrivastava, S. and Mitra, S. (2014) *Mobile Water Kit (MWK): a smartphone compatible low-cost water monitoring system for rapid detection of total coliform and E. coli*. Analytical Methods Vol. 6, No 16, pp. 6236-6246.

Medhi, I., Patnaik, S., Brunskill, E., Gautama, S.N., Thies, W. and Toyama, K. (2011) *Designing mobile interfaces for novice and low-literacy users*. ACM Transactions on Computer-Human Interaction (TOCHI) Vol. 18, No 1, pp. 1-28.

Nickson, A. and Franceys, R. (2003) *Tapping the market: the challenge of institutional reform in the urban water sector*. Palgrave Macmillan: New York, NY.

Nsiah-Gyabaah, K. (2001) *The looming national dilemma of water crisis in peri-urban areas in Ghana*. (DFID Project Funded Project R 7330). DFID: Kumasi, Ghana.

Ochara, N.M. and Mawela, T. (2013) *Enabling social sustainability of e-participation through mobile technology*. Information Technology for Development No, pp. 1-24.

Siedner, M.J., Lankowski, A., Musinga, D., Jackson, J., Muzoora, C., Hunt, P.W., Martin, J.N., Bangsberg, D.R. and Haberer, J.E. (2012) *Optimizing network connectivity for mobile health technologies in sub-Saharan Africa*. PLoS ONE Vol. 7, No 9, pp. e45643.

Stoler, J., Weeks, J.R. and Appiah Otoo, R. (2013) *Drinking water in transition: A multilevel cross-sectional analysis of sachet water consumption in Accra*. PLoS ONE Vol. 8, No 6, pp. e67257.

WHO/UNICEF. (2011) *Drinking water: Equity, safety and sustainability*. World Health Organization: Geneva.

WHO/UNICEF. (2014) *Progress on drinking water and sanitation, 2014 update*. World Health Organization: Geneva.

**Contact details**

Justin Stoler
University of Miami
Department of Geography and Regional Studies
Coral Gables, Florida, USA
Tel: +1 305 284 6692
Email: stoler@miami.edu

Clay Ewing
University of Miami
Department of Cinema and Interactive Media
Coral Gables, Florida, USA
Tel: +1 305 284 2995
Email: c.ewing@miami.edu