Enhancing Water Service Sustainability by leveraging digital public goods and bridging the trust deficit with Water infrastructure data

Sunderrajan Krishnan and Kiran Kumar Sen

Abstract

The challenge of delivering and sustaining improved water services in India is significant due to the distance between district centers, Gram Panchayats, and individual homes, leading to inequities and failures in public services. The Jal Jeevan Mission (JJM) exemplifies rapid infrastructure investment but highlights the need for ongoing government-community interactions to maintain these services. Traditional digital governance models are top-down, excluding civil society actors from design and delivery. However, new open-source technologies enable organizations like INREM to facilitate digital interfaces between communities and governments, enhancing water supply sustainability.

Greater transparency in governance is leading to innovative policy designs. The JJM, for instance, has introduced a Citizen Corner to boost citizen engagement. Assam's JJM is a notable example, empowering village communities to address water supply problems independently through increased communication with government officials.

Deploying human-like conversational abilities at scale can increase interactions with field functionaries of water supply programs. A WhatsApp-based bot is being tested to address water contamination issues. However, these trials revealed governance challenges. Community agency is most enhanced when actionable, real-time suggestions are provided. Yet, fact-checking requires slow and expensive human involvement, reducing user interest.

Given the sensitive nature of data shared in conversational interactions, data privacy and secure data storage/access are critical concerns. Deploying complex digital systems for human-like conversation-based problem solving in water and environmental issues is still in its infancy. It's essential to consider basic, reusable building blocks available as open digital assets to build these solutions effectively.

This streamlined approach can foster stronger government-community interactions, leading to more sustainable water services.

Keywords: Drinking Water Policy, Water Quality, Public Participation, Emerging Technologies, Digital Public Goods

Introduction

Rural India mostly drinks groundwater with almost 85% of population being groundwater dependant for their drinking water needs (TERI, 2022). There are exceptions now across Telangana and large parts of Gujarat where regional pipe water supply schemes deliver treated surface water to villages. This water is supplied mostly from local deep groundwater sources. People drink either this supplied water, or water from other groundwater sources near their homes, or springs. There is also an increasing trend, even in rural India, of buying drinking water from water service providers.

As a consequence of government programmes, education, market push and diffusion of information in general, rural communities are now more conscious of what kind of water gets supplied. In general, there is revulsion for water that is muddy, smelly or anything that tastes really bad. These mostly constitute what are known as the Physical Parameters within what is the Drinking Water Quality Standards ratified by the Bureau of Indian Standards (BIS). There are then a series of Chemical parameters and a lot more totally around 35 such substances that need to be within limits.

Only a small fraction of rural India is conscious about the Drinking Water Quality Standards as prescribed by the BIS. The Central government and State governments have taken up the responsibility to deliver Safe drinking water till people's homes. However, the ability to convert Water Supply Infrastructure into reliable, safe and trustable systems is a journey yet to be travelled.

The problem of monitoring Quality of supplied water, involving people in it, and bringing more trust has been the focus of multiple policy efforts since the past 2 decades. These have also addressed the possibilities of people making their own choices and relying on other water sources - with monitoring of groundwater broadly also being done. We first try to understand here the usefulness of such efforts from a lens of Scale, People Participation and Perception of accuracy - and then look at policy strategies that can help bring more trust into the quality of water being supplied.

Water Quality Monitoring Systems

The Year 2005 saw a big push by the Government of India towards people's participation in generating Water quality data across India (DDWS, 2005). Publication of the uniform protocol for Water quality monitoring and surveillance (WQM&S), later revised in 2013, have now been consolidated with the Jal Jeevan Mission (JJM) and given much wider visibility with the platform for WQ-MIS (Water quality management information system). This progress across the past 17 years is now bearing fruit with more than 1.5 million sample points emerging on Water quality in the year 2021-22 from 82,000 villages across India (as of Jan 29, 2022). Some states have emerged as champions in this process, such as Chhattisgarh, which alone contributes 6.5 lakh sample points. Quietly and surely, a revolution is in process and this will have its own impactful repercussions over the next few years. As 5 women in every village are trained to test and report Water quality data in a public digital platform, it is rightfully bringing much optimism towards Functional Household Tap Connections (FHTC) and the allure of Safe water within homes.

The rise of participatory WQM&S within villages and its communication to people, with action being taken through Village Action Plans (VAP), brings a much needed foundational structure to the entire system of Water quality monitoring being done by different agencies. Now, therefore, is the best time to open up the layers right from the villages, and look at how other data generation systems on Water quality, their respective monitoring networks, and their respective communication systems are working. We will spend the next section trying to understand the current architecture, and then come back again, to the core questions, of the compatibility and joint usage of these systems towards communication and action.

Architecture of water quality data collection

A number of institutions monitor water quality, and especially, we focus here on Groundwater, and related measurements, which have a bearing on different uses[1]. Some of the institutions monitoring groundwater quality are CPCB, CGWB, PHED Laboratories, ICAR, Ministry of health, JJM WQM&S, Indian university groups, Research institutions, NGOs, Citizen groups, International universities and INGOs. Increasingly, satellite based estimation, newer indirect measurements with machine learning, and other means are being attempted, not only for surface water, but also for groundwater.

These water quality data systems vary in terms of a) what they measure, b) how often they measure, and c) at what scale or density their data points are spread.

Primarily, most measurements are focussed on basic indicators of water quality such as Physical parameters (turbidity, colour and others), pH, TDS, Electrical conductivity, and on some of the commonly occurring major and minor ions (with focus on fluoride, arsenic, iron, nitrate). Based on interest, such as from health ministry, there is emphasis on pathogen indicators such as coliforms, but very less on more comprehensive microbiological testing. Beyond this, measurement of pesticides, VOCs, SVOCs, PAHs, and other parameters are very few, For example, for radiologically relevant parameter such as Uranium, there has been a specific national project coordinated by the Department of atomic energy in participation with research institutions across the country.

Frequency of measurements, are at maximum twice a year, for regular monitoring of the basic parameters, and in some cases of special interest and focussed research projects could be in much finer frequency of monthly or even lesser. For the lesser measured parameters such as pesticides, and radiological parameters, measurement is based on need and could be a one time process in most cases. There is a wide variability of frequency of measurement across data generation systems and parameters.

The scale or density of measurement ie number of monitoring stations, varies from a range of 1 data point per district or block in the case of CPCB/CGWB, to one per GP scale (PHED labs) and one per habitation scale (JJM WQMIS). The density variation is therefore more than 1:1000 variation across these data monitoring systems, and that has a bearing on the quantum of data available from these different data sources.

Scales and density, participation and perceptions

The purposes for which Water quality data needs to be utilized are highly varied - Water resource management/river basin planning, Hydrogeological investigations, Irrigation, Drinking water supply, Public health management (controlling epidemics and chronic

[1] This section and the next has been supported by communication with Peter Ravenscroft, as part of a consultation on Water quality data governance, along with World Bank in India

problems), Environmental management (streams, rivers, lakes, near ocean), Industry (for both supply of water to industries and management of effluents from industries), basic fundamental research, Environmental impact assessments.

Based on the purpose, there is need for specific Water quality data of those respective parameters, need for accuracy, and of scale and density with localization. We take examples of 3 water quality data generation systems from all of the above to discuss aspects of scale/density, participation and perceptions with respect to usage of the data.

CGWB: The focus of CGWB being that on groundwater as a resource, is specific to understanding the status of how aquifers are in terms of their exploitation as quantity and quality. There are approximately 15,000 monitoring wells across India. Basic major and minor ions are monitored twice a year across these stations, so for this discussion, we can take 2 x 15,000 = 30,000 and broadly 50,000 as a scale of measurement of sample points annually. Apart from these, CGWB conducts in depth studies of aquifers going into specific local geographical distribution of water quality parameters. The NAQUIM initiative anchored by CGWB has been able to produce specific local maps of Water quality in areas of focus across the country. Majority of focus of CGWB monitoring wells are that of shallow aquifers, with a specific focus on deeper aquifers where needed.

PHED Labs: There are 2020 active labs of the PHED across India (as of January 29, 2022). As per yearly plans of the district level labs, there is an intention of 3000 sample point measurements every year of drinking water sources within each district, with twice a year measurement. Though commonly, the basic major/minor ions and around 14 water quality parameters are measured for each sample, there is a possibility based on local infrastructure/capacity to specifically measure other parameters when needed. As of the year 2021-22 , 30 lakh sample points have been measured in these labs (as of data January 29, 2022).

JJM WQMIS: The aim of WQM&S programme is to equip more than 6.4 lakh villages of India to be able to test for water and report it with the WQ-MIS portal. The field testing kits (FTK) that are supplied to villages are used by a variety of frontline workers and volunteers, based on the specific state. These FTKs have the capacity of 20-100 samples points every year, so potentially around say, 50*6 = 30 million measurements are possible when this programme is in full operation. As of the year 2021-22, it already has scaled up to 18.6 lakh measurements from 80,000 villages (dated as of January 29, 2022). The basic focus of FTK is to be able to test the primary water quality parameters, mainly 5-6, but potentially 10-14 parameters that are possible using these kits.

	CGWB	**PHED Labs**	**JJM WQM&S**
Scale (potential)	x = 50,000 samples a year	50 x	100x-500x
Participation	Low to minimum	Samples contribution by	Participation by

		community in some states such as WB	some community members
Perception of reliability	High and supported by specific studies, of localized nature and supported by NAQUIM aquifer mapping programme	Increasing reliability with NABL certification, but possible inadequacies existing due to infrastructure and transparency issues	Low level of trust on reliability amongst technical staff and scientists, especially because of indicative nature of FTKs and possible inaccuracies in sample collection, testing and reporting

We see here that measurement of the problem for planning public policy on water supplies is dependent on multiple sources of data. Either we have a problem of scale, or a problem of trust. If locally relevant and dynamic data is to be made available, there is a trust issue as we have today.

The Nature of the problem as it is today

The programme of people's participation in Water Quality Monitoring and Surveillance (WQMS) as described above has been seen as one approach to support water supply programmes to deliver safe water to households. This is seen as one of the flagship approaches to the Jal Jeevan Mission (JJM) programme.

However, this programme of WQMS fails due to the perspectives from different actors who are involved in it.

A Community member perspective: Volunteers are 5 women from every village, who need not be recognized frontline workers of any department. Some minor incentive is paid, but is not enough even for travel expenses. A strong perspective is that this task is not worth the time and money. Secondly, entering data into WQMS digital systems is a cumbersome process and not designed for community level users. Even if this data shows a problem and it is submitted there is low trust on any action happening based on the data.

A Water engineer perspective: Too much of the community generated data adds to the pressure in supplying water. The community also needs to take responsibility and ownership, otherwise, it will only add to the engineer's work and burden.

A Lab chemist perspective: There is low trust in the quality of information being shared by the community. The 5-women volunteers are not seen as being capable to test for water quality and doubt is cast over the uncertainty of this information.

Can we bring more trust into this relationship between community members, water engineers and chemists, who are important actors in the WQMS system within any district. How do we let the engineers see value in WQMS data, How do we get the chemists to trust this data more, and How do we get community members to also take their own responsibilities and expect that engineers will respond based on their contributed data.

What can we learn from how these problems have been approached elsewhere

Right from large scale population level architecture to localized systems, different approaches are being attempted in other sectors such as education, health, poverty alleviation and climate, which can help us here.

Migration: Bandhu is a means of helping migrant workers search for affordable rental housing in cities. Here, it addresses the problem of information asymmetry which housing agents exploit, at the loss of migrants. The information and trust gap is filled by using open source tools and databases that are brought together with very less knowledge of technology.

Education: Shiksha Lokam addresses leadership in school education by helping small improvements made possible through assessments made themselves by teachers. Here, an open micro-services architecture technology helps not only this mission, but produces numerous small tools that are potentially of help to many other missions.

Climate: Samaaja helps volunteer engagement spread out geographically to be recognized for their localized problem solving on climate issues. Reap Benefit uses it for youth engagement, but by offering it openly, the same system is used for solving water problems, in health and other issues, all of which are trying to bridge the gap between community representatives and larger systems such as government programmes.

These examples above tell us about the rising missions in which problems are being looked at in a generic manner, and technology is either repurposed of freshly created with an intention of being usable for other missions from totally different sectors.

Some experiments and newer challenges

INREM has been working closely with the Jal Jeevan Mission (JJM) to see how some of the problems of WQMS can be solved and experiences from other sectors be utilized here.

We defined three specific areas where technology can act as an enabler:

- Making data accessible
- Contributing data
- Supporting individual actions

Public data contributed to the JJM is made available through an interface of a WhatsApp bot. Data can be accessed through this *OurWater* bot by sharing the user location. Users can query with text, voice messages, and images, with responses being guided by AI through the bot.

The technology architecture here is by using microservices rendered from solutions which are open access or Digital public goods (DPG). We utilise Glific for conversational bot interface, Dalgo and Google cloud for data storage with an API that helps to 'Get' data. APIs from Glific to OpenAI supports interpretation of user queries, based on a knowledge-base that is open sourced.

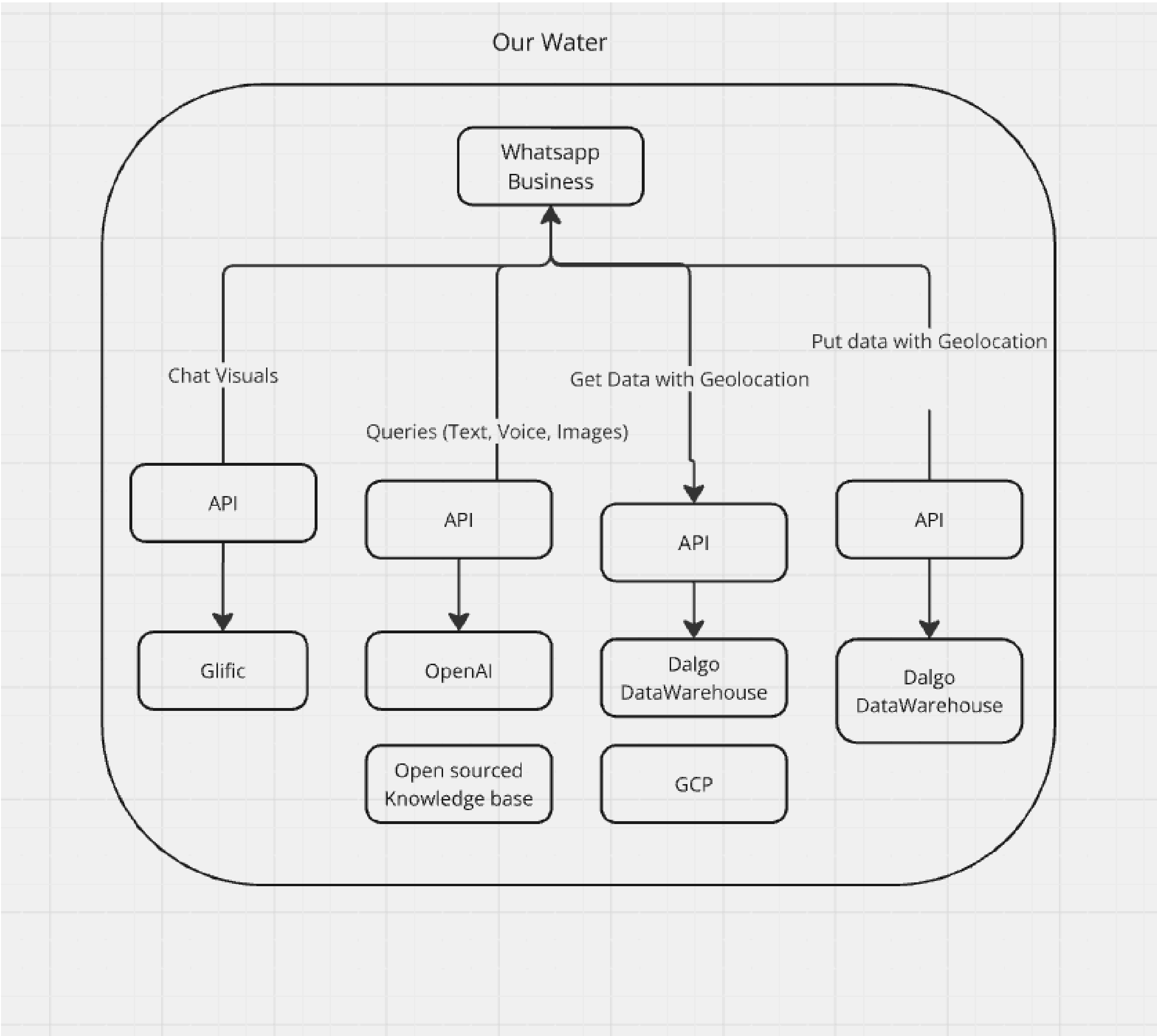

Figure: Schematic diagram of microservices architechture for OurWater bot

As a whole, the interface is built such that each component can be developed independently and any other open source solution or microservice can be brought in to fill the same need. Such a design helps scaling with larger public systems making customization and special needs possible.

Learning from experiments

The OurWater bot described above has been distributed for a first level user test, resulting in= approximately 850 user registrations across 8 states. These users have interacted with the Bot in their respective languages of comfort as shown below in the graph. Due to the nature of locations, we see most interactions in Hindi here.

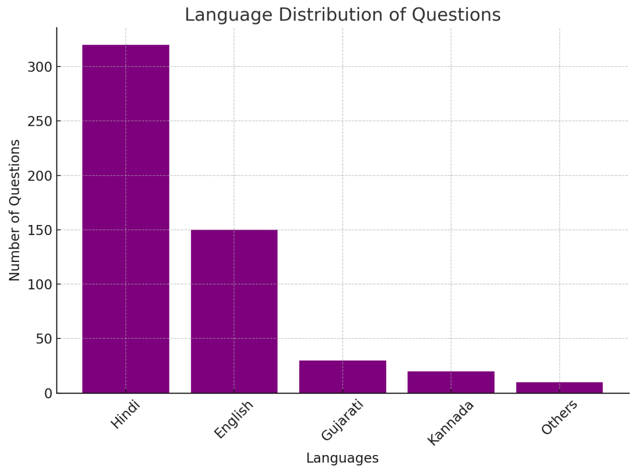

Figure: Language distribution of queries on the OurWater bot

In terms of language comfort, we found that automated translation results into a bit of formal language structure, and it might be necessary to bring a more informal interaction here when it comes to translation. Users then interacted on different aspects ranging from accessing data, to contributing data as well as asking queries of their concern. We try to glean some patterns coming out from 500+ user queries in the system.

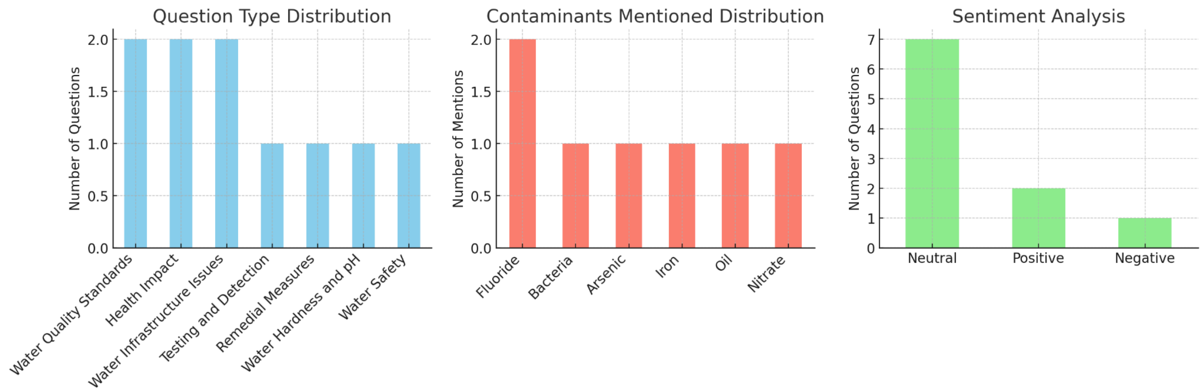

Figure: Summary of queries posted on the OurWater bot

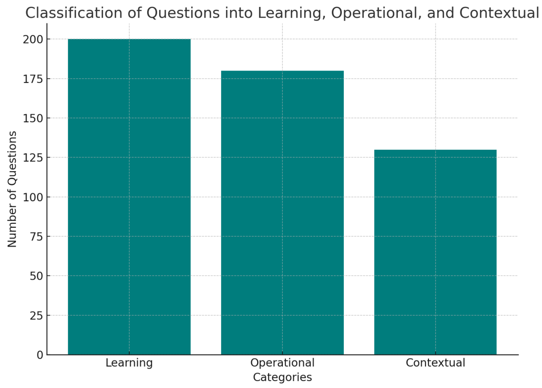

Figure: Classifying queries into the nature of inquiry

The nature of queries on the OurWater bot range from questions clarifying for better learning, to more contextual queries specific to the particular locations where people stay. For example: We have questions like: How much Nitrate is bad for me? , and some others like, In my place, I have oily water. What should I do?

Our understanding of these questions also shows that there is a relatively larger neutral tone to the questions asked here, but a little bit of positive or negative emotion. This tell us that engagement till now is not very emotional, and possibly given the nature of a technology test.

Further, we have also attempted to take out a summary of what people ask for, in terms of an FAQ which is summarised below.

1. What is the permissible limit of fluoride in drinking water?
2. How do I test for pathogens in my drinking water?
3. What are the health effects of drinking water with excess fluoride?
4. How can I remove fluoride from my drinking water?
5. What are the symptoms of arsenic exposure from drinking water?
6. What is the recommended pH level of drinking water?
7. What is 'blue baby syndrome' and how is it related to nitrate in water?
8. How do I know if there is too much iron in my drinking water?
9. What is the acceptable limit for total dissolved solids (TDS) in water?
10. How can I ensure my water is safe to drink?

Figure: FAQ derived from the user queries on OurWater

This initial analysis tells us that, there seems to be a general curiosity with querying related to what people experience and an assistance such as what is being provided here could be of use to the type of users encountered here. We also summarize below the contributed data (around 800+) across the country.

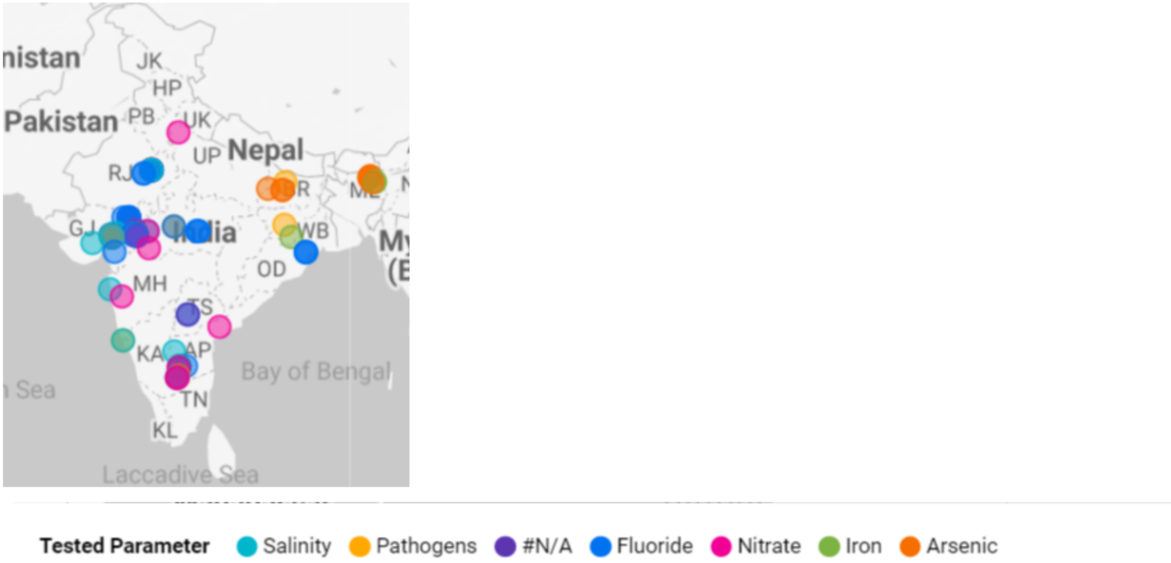

Figure: Map of 800+ user contributed data from the OurWater bot

All of the above raise questions for us, as below:

1. Trust with contributed data and verification: While we can see that the system built here is able to engage the rural community audience and make it easy to contribute data, the element of trust and verification needs to be looked at.
2. Automated responses: With respect to AI generated responses, even though constrained by a knowledge base, we have issues of authenticity and how the user then responds with action to such responses. Here ethical issues arise and need to be understood better.
3. Lastly, very importantly, the consequence of such interactions, as in, what happens next, how local problems are solved with more participation, is an important area of work and much development needed on that front.

Summary

In this paper, we look at the general problem of trust deficit and gap between public data and accountability, specific to Water supply in rural areas. When we take data currently available on Water quality, there are a variety of sources, with varying resolution and perceptions of trust. While community participation is a response in policy and programmes, it ails from poor participation and trust. We have experimented here with learning from other sectors on reusing open technology and building a microservices based architecture called the OurWater bot. Early stage tests of this bot show some leading answers to the participation side of this problem, but also give us challenges to consider and do more work on the trust side of the problem.

Water service from programmes such as the Jal Jeevan Mission will only sustain if there is continued ownership and participation from community members. Experiments with new and emerging technologies is showing one path, but we also see caution here, keeping in mind the social inequities and complexities that are associated with water.

References

TERI, 2022,Water Crisis in India: The World's Largest Groundwater User, https://www.teriin.org/article/water-crisis-india-worlds-largest-groundwater-user

DDWS, 2005, Uniform Drinking Water Quality Monitoring Protocol