

# **The Quality of Civic Data in India And the Implications on the Push for Open Data**

Satyarupa Shekhar and Vinaya Padmanabhan  
Transparent Chennai, IFMR

## Executive Summary

Data is valuable because it has the power to transform economic, political and social processes. Businesses can develop codes and tools based on the data. Citizens use these data derivatives to access information about the weather or navigation. They also use government data to monitor public services. Governments also use open government data; in some cases, they are often the most proliferate users of open government data. This multiplicity of benefits has lent strength to the demand for open data. Open data has the potential to make government more transparent, accountable, and more democratic by enabling greater public participation. For groups representing the poor and disempowered, it has the potential to enable them to hold governments accountable for their needs. For others, open data has the potential to make governments more efficient through data-driven decision making.

Openness is being driven internationally through multilateral associations such as the Open Government Partnership. It can also be implemented at three levels in a country: at the federal, state, and city levels. There has been progress at all of these levels in India, the most ferment being observed at the city level, perhaps because cities are the first point of contact with citizens but also because there has been so little data that city governments have made open. Cities are also the focus of urban renewal projects and investments that are premised on their being data-driven in their decision making.

Our local experience suggests that the push for open data as it is currently framed nationally and internationally assumes a far higher quality and comprehensiveness of existing government data at the city level than what currently exists in Chennai. Our research sought to investigate the quality of municipal data and to demonstrate that the implications of poor quality data on civic services are enmeshed in larger institutional and organisation contexts of that civic service. We examined data availability, quality, processes of creation and use of data, and impact of data quality using a methodology that used unstructured and semi-structured interviews, surveys, physical and digital mapping methodologies, public consultations and focus group discussions to answer the following questions:

- Does the evidence collected suggest that government data quality is indeed questionable?
- If so, what are the ways in which this data is deficient, particularly with regards to the poor?
- Do these deficiencies result from common failures in data collection and storage processes?
- Does this poor quality of data actually impact the quality of services provided to residents, particularly the urban poor?
- Finally, are there ways in which this data can be easily corrected or improved?

We found that there are several problems with the data available with municipal authorities, including incomplete and inaccurate data, a lack of data standards, formats, metadata and licenses, no integrated data repositories, selective public disclosure and negligible sharing of data between government agencies. This directly results in poor planning and

maintenance of public infrastructure, and poor provision of public services, particularly to the urban poor. Poor data sharing practices reinforce the fragmentation in governance that already hinders public service provision. We also find that poor data and sharing practices impedes transparency and accountability. Opening up data can help overcome the obstacle of sharing between public agencies and may also improve internal accountability. However, we caution against making existing data open without first addressing the causes of poor data quality has the danger of reinforcing existing social, economic and political divides. We found that the city government was willing to try new methodologies to create data to fill gaps but efforts need to be made to institutionalise these processes.

## Table of contents

Glossary .....	5
1. Introduction.....	6
Key Research Questions .....	8
Methodology .....	9
Institutional Context.....	11
2. Research findings.....	13
Inaccurate data .....	14
Incomplete data .....	15
Lack of data formats and standards .....	16
Selective disclosure.....	17
Lack of metadata and licenses .....	18
Negligible sharing of data.....	19
No integrated repositories of data .....	20
3. Discussion .....	21
4. Conclusion .....	25

The report was authored by Satyarupa Shekhar and Vinaya Padmanabhan, We acknowledge the research support extended by other Transparent Chennai researchers Aishwarya Balasubramaniam, Dhivya Murugaian, Gulika Reddy, Kadambari Badami, Nithya Raman, Prabu Raja, Parthipan Duraisamy, Priti Narayan, Ranjeet Joseph, and interns Priya Govindaraj and Tryphena Duddley.

This case study research was done under the Exploring the Emerging Impacts of Open Data in Developing Countries (ODDC) research project, funded by the World Wide Web Foundation supported by grant 107075 from Canada's International Development Research Centre. For more see <http://www.opendataresearch.org/emergingimpacts> 2014.

This work is licensed under a Creative Commons Attribution 4.0 International License.

## Glossary

ABBREVIATION	
AC/ JC/ DC	Additional/ Joint/ Deputy Commissioner, always an officer of the Indian Administrative Service (IAS)
AE/ JE	Additional / Junior Engineer, is responsible for infrastructure and services at the ward
AEE	Assistant Executive Engineer, typically supervises AE/JEs at the unit (a group of wards)
CE	Chief Engineers
CMWSSB	Chennai Metropolitan Water Supply and Sewerage Board, a para-statal agency. It is administratively part of the MAWS department.
CoC	Corporation of Chennai, the city government of Chennai. It is administratively part of the MAWS department.
EDP	Electronic Data Processing, the IT department of the CoC
EE	Executive Engineer
ERP	Enterprise Resource Planning
GIS	Geographical Information System
IAS	Indian Administrative Service
MAWS	Municipal Administration and Water Supply department
NDSAP	National Data Sharing and Accessibility Policy
PGR	Public Grievance Redressal
PS	Principal Secretary, an IAS officer
RAC/ RJC/ RDC	Regional Additional/ Joint/ Deputy Commissioner, always an IAS officer
RTI	Right to Information
SE	Superintendent Engineer
SWD	Storm Water Drains
SWM	Solid Waste Management
TC	Transparent Chennai
TNPCB	Tamil Nadu Pollution Control Board
TNSCB	Tamil Nadu Slum Clearance Board
ZO	Zone Officer; a zone is a group of wards

# 1. Introduction

Open data has support from across the ideological spectrum because it holds a bouquet of promises. It has the potential to make government more transparent, accountable, and more democratic by enabling greater public participation. For groups representing the poor and disempowered, it has the potential to enable them to hold governments accountable for their needs. For others, the focus is on good governance: open data has the potential to make government more efficient through data-driven decision making. It also has the potential to make government data available for commercial uses, like transportation apps and mapping services, that could both create profit for companies and make government information far more useful for residents. As a result, the open-data movement has seen some unlikely bedfellows, from activists and community groups, to for-profit corporations, non-profit organizations and donor agencies, and even governments themselves rallying to open government data, to use technology to promote its use, and to enshrine reforms about open data in national and sub-national laws.

Openness can be driven at three levels in a country: at the federal, state, and city levels. There has been progress at all of these levels in India. At the national level, India already has a Right to Information Act, passed in 2005, which gave citizens of the country the constitutional right to access government data through an application process. In recent times the Indian government has sought to align its policy and practice with a global trend by countries, such as the USA and UK, by actively encouraging open data. The Indian government recently approved the National Data Sharing and Accessibility (NDSAP), a policy that promotes proactive disclosure of data by the government.<sup>1</sup> It also launched data.gov.in in 2011, an open data portal that publishes government data from different departments. At the sub-national level, state governments in India, too, have passed their own RTI laws in accordance with the national law. However, there are no city level policies that mandate proactive disclosure. Tamil Nadu, the state where Chennai is located, has no central data repository and very little sharing of data even between government departments.

The level at which there has been the greatest amount of ferment around the question of open data is that of the city. Perhaps this is because city governments are the first point of delivery of public services. It could also be that despite being the first point of contact with citizens, few city governments have disclosed high quality data. City governments have taken steps towards improving databases and maps for planning, sometimes driven by local civil society organisations and hackathons organised by interested individuals, as well as supported by central government funding for specific projects such as improving GIS databases. City level reforms in data and technology use have also been hastened by the central government, especially since a new government came to power in May 2014, which has focused on smart cities. In July 2014, the newly elected national government stated that it was going to allocate Rs. 7,060 crore or approximately USD 1.2 billion to make 100

---

<sup>1</sup> Department of Science and Technology, “Introduction to the National Data Sharing and Accessibility Policy (NDSAP)”, accessed May 29 2014, <http://www.dst.gov.in/nsdi.html>

smart cities in the country.<sup>2</sup> According to the Ministry of Urban Development, Government of India, a smart city is one that primarily provides employment opportunities to its residents.<sup>3</sup> The Smart Cities Council, an advisor to the Government of India on its Smart Cities Mission, defines a smart city as one that embodies livability, workability and sustainability and "... gathers data from **smart devices and sensors** embedded in its roadways, power grids, buildings and other assets. It shares that data via a **smart communications system** that is typically a combination of wired and wireless. It then uses **smart software** to create valuable **information and digitally enhanced services**".<sup>4</sup>

The central government's approach seems to be one that has strong corporate leadership. A link from the Ministry of Urban Development's website takes us to the Indian Smart Cities website, where the definitions of a 'smart city' have been proposed by IBM, CISCO, Frost and Sullivan and the UK Department of Business, Innovations and Skills.<sup>5</sup> There is also an industry body, the Smart Cities Council that has Alstom, CISCO, Bechtel, GE, IBM, MasterCard and Microsoft as the lead partners, and has cited reports that estimate the market for smart cities to be USD 8 billion today and expected to be USD 2 trillion by 2020.

It is in this context of strong global and national lobbying and advocacy for policies and investments to make Indian cities 'smart cities', and parallel, though not always integrated, movements for open data, that our research study examined the quality of civic data in the Chennai. While the presence of corporate support for smart cities is not in itself problematic, our reading of these documents suggests that this push towards smart cities has been focused on technology rather than on the potential gains for citizens in terms of improved transparency, accountability, and responsiveness. Our research stems from our interest in ensuring that open data truly benefits residents, including the urban poor.

Our local experience suggests that the push for open data as it is currently framed nationally and internationally assumes a far higher quality and comprehensiveness of existing government data at the city level than what currently exists in Chennai. This also raises challenging issues for UN calls for a Data Revolution. Data is often outdated and incorrect or incomplete. Data is often not accessible to the public, and even when available is often in formats like paper or PDF that are not easily used for sharing, analysis, and monitoring. Even more problematically, it is not shared across departments, even when multiple departments require the information for efficient completion of their tasks. Though there is a large amount of data about the city, spatial and non-spatial, digitised and otherwise, that is available with different government departments and agencies, this data is collected and stored in various formats and locations, making it difficult for both citizens and public officials to access it. There is no central data repository that can facilitate sharing and consequently, different agencies either do not have complete data or duplicate that collected by other agencies. Government officials admit to being overwhelmed by RTI requests for government data and their inability to provide comprehensive and correct

---

<sup>2</sup><http://indiatoday.intoday.in/story/pm-narendra-modi-100-smart-cities-key-elements/1/382021.html>

<sup>3</sup>[http://indiansmartcities.in/downloads/CONCEPT\\_NOTE\\_LATEST\\_25914.pdf](http://indiansmartcities.in/downloads/CONCEPT_NOTE_LATEST_25914.pdf)

<sup>4</sup><http://smartcitiescouncil.com/article/our-vision>

<sup>5</sup><http://indiansmartcities.in>

information. Despite the RTI Act, officials often use their discretion to provide information or to delay indefinitely. In this context of poor quality data, any push for open data will, in parallel, have to include a push to improve the quality of the data as well, and to ensure that the data used for planning in the future does not lead to the further exclusion of the poor simply because of the poor quality of available information.

## **Key Research Questions**

Across civic issues, we have also found that civic data frequently fails to account for the informal sector – neither unrecognised housing settlements nor unregistered economic activities. Unfortunately, this means that most municipal data does not see the poor, because these are the settlements and workplaces where the majority of the urban poor live and work. National statistics, like the Census, collect data on street vending as a source of income, but the city of Chennai does not have this data. Similarly, though the Census 2011 enumerated slums households, the city government relies on a list of slums per ward, which it uses with latent knowledge of the number of households and locations such as, on the banks of a river or lake, or on the seashore.

Local bureaucrats are unaware - except anecdotally - where street vendors operate and what kinds of goods and services they offer. There is also no accurate data about where slum dwellers access water and sanitation facilities, and often they are blamed for sewage outflows to the city's waterways. Not only does this unfairly place them in the path of evictions during river restoration projects, despite documented evidence of large residences and commercial establishments releasing their sewage into Chennai's waterways. It also fails to notice the impacts on health of the urban poor and misplaces river restoration funds. We believe that these omissions are important, and impact the daily lives of the poor – and other city residents – in important ways.

As a result, our experiences with civic data suggest that opening up this data alone may not be enough to actually bring about greater government transparency and accountability at the city level. Indeed, relying on existing data for better urban management may actually lead to the marginalisation of communities that are already vulnerable. In Chennai, before open data can lead to improvements in the city, there needs to be a closer examination of the data itself.

The project examined data availability, quality, processes of creation and use of data, and impact of data quality. The case studies will help us to answer a number of related questions:

- Does the evidence collected suggest that government data quality is indeed questionable?
- If so, what are the ways in which this data is deficient, particularly with regards to the poor?
- Do these deficiencies result from common failures in data collection and storage processes?



- Does this poor quality of data actually impact the quality of services provided to residents, particularly the urban poor?
- Finally, are there ways in which this data can be easily corrected or improved?

By identifying common problems with data quality, we hope to develop common methodologies and/or general principles through which Indian and other developing country cities can better account for the informal and the urban poor in their existing numbers.

This research project was designed to help us, an action research group external to government, develop a policy platform around open data that responds better to our local context. The project allowed for a more nuanced understanding of the existing municipal data, to document the gaps in the data and how they come about through flawed or incomplete data collection processes, understand the impacts of these data gaps, and how these gaps can be filled either through improving existing processes or by allowing an opportunity for citizens to interact with the data.

This research enabled us to advocate for a nuanced approach to collecting and using data, one that emphasises the limitations of existing municipal data, pushes for improvements in existing processes, and provides for means of contesting and improving existing data, not just using it. We were able to establish a two year Memorandum of Understanding with the city government to bring this learning to bear on its data management practices. As part of this collaboration, we worked with city engineers to define performance indicators for various municipal infrastructures (parks, footpaths, toilets, bus shelters, construction and demolition debris, etc) and identified metrics to measure performance. We also developed an outcomes based tender that incorporated these indicators and linked them to payments of contractors hired to maintain these infrastructures. However, we needed to develop these frameworks in more detail and build the capacity of the engineers to monitor performance contracts. Unfortunately, we faced resistance from certain groups of engineers who were uncomfortable with the internal accountability that our processes demanded, and we terminated our agreement ahead of schedule.

We expect that through the research and experience with the city government of Chennai, we will be able to uncover strategies that communities could use to respond to issues of poor data, by correcting faulty records or lobbying against them or by changing the way in which these datasets are used or interpreted, not just in Chennai but in other Indian cities as well.

## **Methodology**

The research aimed to assess three different kinds of municipal data - data on access to water, public sanitation, and public health – to see whether our supposition about the poor quality of this data held true across these three sectors. These sectors were selected because they are all data about the city that were likely to be available under future open data initiatives, and they are very important for the city's poor.

We began the research project from April 2013 through September 2014. Since our collaboration with the CoC overlapped - from February 2014 through May 2015 - we were able to gain access to officials and gather insights that had not been part of the original scope of the project. We have included brief references to these while retaining our original focus on water, sanitation and health.

We used a combination of methods to gather data and information about the state of civic data in Chennai. We used unstructured and semi-structured interviews, surveys, physical and digital mapping methodologies, public consultations and focus group discussions.

For each issue, we attempted to answer the following questions, which are divided thematically:

#### *Stocktaking*

- 1) What kind of data is collected on these issues by the municipal and state governments?
- 2) Which agencies collect this data? Where is the data currently stored?
- 3) How easily is the data available to the public and to other government agencies? What kinds of formats is the data currently stored in?

#### *Understanding the process*

- 4) What are the processes through which this data is collected? Which agencies, offices, and individuals are involved in data collection? How frequently is the data collected? How is it verified (if at all)?
- 5) How is the data currently used in planning, policymaking, budgeting, and monitoring (if at all)?

#### *Evaluating data quality*

- 6) Does this process result in data that accurately reflects conditions on the ground? Can the gap between official statistics and the real incidence of the phenomenon described by the data be quantified or estimated in any way?

#### *Understanding the real-world impacts*

- 7) What are the implications of poor data quality on government performance in each of these sectors city-wide?
- 8) What are the impacts of the inaccuracies of this data on local communities? Can these impacts be demonstrated or documented in a single neighborhood or individual?

#### *Developing a response*

- 9) If the data is not accurate, what kinds of changes in the data collection process would make the data much more accurate?
- 10) Are there examples at the city-wide level or at the neighborhood or individual level where residents or city employees have successfully worked to bring to light, supplement, or correct erroneous data in these sectors?

This report presents our research findings in an attempt to answer these questions, while the individual case studies provide a sector-specific understanding of the context, problems and implications of poor quality data. What this report also aims to do is demonstrate that the myriad problems and implications of poor quality data on civic services are enmeshed in the larger institutional and organisation contexts of that civic service. These are, in turn, being affected by the national and global policy environment which is influenced by commercial interests and corporate agendas. This report also presents an example of an ICT-linked urban innovation that was the result of a collaboration between a research group and the city government as one alternative to the state-corporate entangled digital agenda.

## Institutional Context

Chennai's local government, the Corporation of Chennai (CoC) was created under the Chennai City Municipal Corporation Act, 1919.<sup>6</sup> Its functions include providing for infrastructure and promoting health and education for the citizenry. The CoC operates within administrative boundaries called zones and wards. The city expanded in 2011 to include the area of 42 local bodies within its jurisdiction (Figure 1). After the expansion the city of Chennai was reorganized into 15 zones and 200 wards (Figure 2). A ward is the smallest electoral unit in the city and has a population of approximately 30,000. A group of approximately 15 wards is called a zone.

	WATER	TOILETS	HEALTH
Administrative area	Municipal Administration and Water Supply (MAWS)	Municipal Administration and Water Supply (MAWS)	Municipal Administration and Water Supply (MAWS)
Agency	Chennai Metropolitan Water Supply and Sewerage Board (CMWSSB)	Corporation of Chennai (CoC)	Corporation of Chennai (CoC)
Organisational type	Parastatal entity	City government	City government
Department	-	Public works - Buildings	Public Health

**Table 1: Relevant Institutions and Organisational Types**

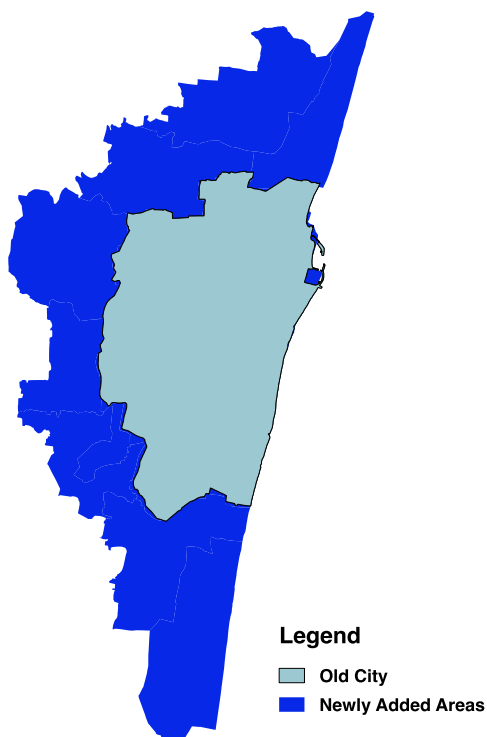
Chennai's local government, the Corporation of Chennai (CoC) was created under the Chennai City Municipal Corporation Act, 1919.<sup>7</sup> Its functions include providing for infrastructure and promoting health and education for the citizenry. The functions of the CoC include providing infrastructure and promoting health and education among the city's residents. The CoC operates within a framework of administrative boundaries called zones and wards. A ward is the smallest electoral unit in the city; a zone is a group of wards and

<sup>6</sup>The Chennai Metropolitan Water Supply and Sewerage (Second Amendment) Act, 1997

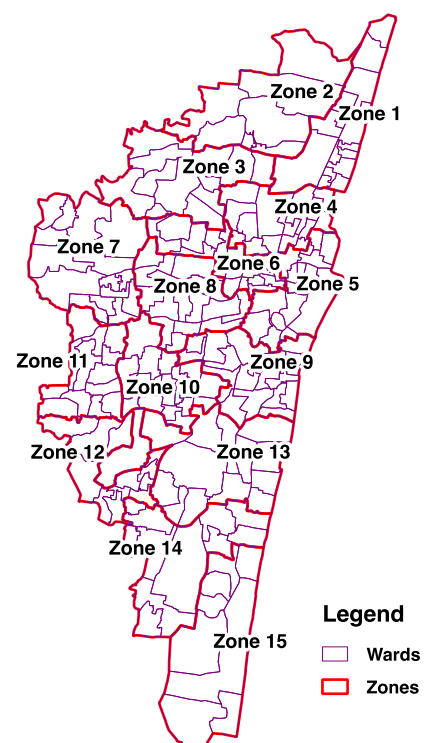
<sup>7</sup>The Chennai Metropolitan Water Supply and Sewerage (Second Amendment) Act, 1997

is an administrative unit. To better manage the rapid development of peri-urban areas, the city expanded in 2011 to include the area of 42 local bodies on its peripheries within its jurisdiction (Figure 1). After the expansion, the city of Chennai was reorganized into 15 zones and 200 wards (Figure 2). Each ward has a ward office where engineers make budgets for new infrastructure requirements and monitor construction. Each zone has a zone office staffed by officers who monitor infrastructure, garbage collection, health, education, and revenue collection at the zone level.

**FIGURE 1: MAP SHOWING CITY EXPANSION**



**FIGURE 2: ADMINISTRATIVE BOUNDARIES IN CHENNAI**



Even though local government bodies such as the CoC have been empowered to govern cities and provide basic municipal services, state governments have created other statutory bodies and agencies that administer specific functions of local government, especially in many large cities. In Chennai, several of these agencies, such as the Chennai Metropolitan Development Authority, the Metropolitan Transport Corporation, the Tamil Nadu Slum Clearance Board, the Tamil Nadu Housing Board and the Chennai Metropolitan Water Supply and Sewerage Board, are involved in servicing the city.

The CoC has a Council that is headed by a directly elected Mayor and 200 councillors. The Council passes resolutions on civic issues and takes decisions on projects, budgets and allocations of resources in the city. The Council is supported by an administration that is

headed by the Commissioner, an IAS officer appointed by the state government. This wing is responsible for executing the decisions of the Council. The Commissioner oversees the work of 15 departments, including the Buildings department, the Public Health department and the Planning department.

Providing services often requires coordination between departments within the Corporation and between multiple agencies. For example, several departments and government agencies are responsible for ensuring that public toilets are constructed and maintained. The Public Health department of the Corporation is responsible for the cleanliness of the city.<sup>8</sup> The Buildings department constructs the toilet structures.<sup>9</sup> The department is staffed by a Superintending Engineer, three or four Executive Engineers, Assistant Executive Engineers and Assistant Engineers or Junior Engineers. These engineers work out of Ripon Building, the CoC's headquarters in the city and monitor the work of engineers stationed at zone and ward offices. At the zone level, sanitation inspectors report to the zone health officers on public toilets and sanitation in the zone. Zone engineers are responsible for their construction, repair and demolition. Water for toilets is arranged by the Buildings Department, usually by constructing a bore well near the toilet.<sup>10</sup> The toilet is connected to a sewerage system maintained by Chennai CMWSSB. The cleaning and maintenance services are either outsourced or carried out by the Corporation.<sup>11</sup>

## 2. Research findings

Theoretically, an open data policy could support greater access to basic services for residents of Chennai in a number of ways. It could enable greater pressure on the government by providing residents more information about their entitlements. It could enable better planning for poor and underserved communities by the government. It could assist in measuring the impact of policies and schemes designed to benefit the urban poor. It could empower citizens with information that could facilitate comparison between what the government claims about its services and with the actual quality of services being provided by the government. It could enable public agencies to better monitor their performance effectively. However, while moves towards open data are important, the impact of open data initiatives can be affected by the quality of data, the methodologies by which data are collected and shared, and the way that data are used by both government and residents. Our research in Chennai found a number of problems with the quality and storage of existing data, problems which are likely to reduce the positive impact of any open data policies in the city. These problems could be grouped into a few categories, which are listed below and illustrated with examples that we encountered during our research.

---

<sup>8</sup> Corporation of Chennai, "Public Health Department", accessed 7 June 2013, <http://www.chennaicorporation.gov.in/departments/health/publicHealthDept.htm>

<sup>9</sup> Interview with PTNGIG, August 1 2013

<sup>10</sup> Interview with POFRM, January 30 2013

<sup>11</sup> Interview with AEVIDF and HHFAAC, January 30 2013

## **Inaccurate data**

Looking at data obtained from the CoC and CMWSSB, we found inaccuracies in datasets that were often inexplicable. For instance, our own earlier research found that the data on the number of public toilets in the city is extremely inaccurate. In 2010, TC researchers compiled a list of toilets by visiting each zone office and requesting a list from zone officials and obtained a total number of toilets in the city of 517. This was in contrast to the number of 714, which was arrived at by the Right to Information Act requests that TC had filed. Every zone furnished a different number for toilets under the RTI than they had given us in person!

The reason for the variation between the data reported by officers and the data obtained under the RTI is unclear. One possible explanation is that data at the city's headquarters differs dramatically from the data maintained at the zones, and that processes for communicating data between the headquarters and the zones appear to be inconsistent. An official at the ward office revealed that data about public toilets is maintained at each of the 200 ward offices in a notebook or register. The zone office has this data in an excel format, but it is not updated frequently. "Sometimes we are asked to come and update the data on the computer there", he said.<sup>12</sup> Another official said data was updated through discussions at the ward level and then sent to the zone office by email.<sup>13</sup> Data at the zone office is then compiled and sent to the city's headquarters, not on a regular basis, but "only when it is required by senior officials".<sup>14</sup> It is possible that the Public Information Officer stationed at Ripon building who is in-charge of responding to petitions had access to a dataset that had not been updated. In some zones, like 8, 9 and 10, the variation could be explained by the fact that toilets were either constructed or demolished in the time between interacting with officials and filing the RTI petition. However, the RTI data for zone 3 and zone 7 had inexplicably high numbers - 51 and 49 more toilets than reported by zone officers – casting doubt on the quality of this data. The only certainty we had after requesting this information is that whether the total number of toilets in 2010 was 517 or 714, it was hopelessly inadequate for the then population of 4.3 million.<sup>15</sup>

Like the Corporation's public toilet data, Metrowater's data on water connections is also riddled with inconsistencies. One example of inaccurate data received from Metrowater comes from the response to an RTI application filed on October 19, 2012 on the number of subsidised water and sewer connections subsidised for Economically Weaker Sections (EWS). Although Metrowater requires applicants to have a sewer connection before a water connection is approved and constructed, in 16 wards they were more water connections than sewer connections. Officials we spoke to were not able to explain this anomaly in the data.

---

<sup>12</sup> Interview with DHDJCK, 21 August 2013

<sup>13</sup> Interview with UFNDGE, August 8 2013

<sup>14</sup> Interview with DHDJCK, 21 August 2013

<sup>15</sup> Directorate of Census Operations, Tamil Nadu, "Primary Census Abstract – Census 2001", accessed November 16 2013, <http://census2001.tn.nic.in/pca2001.aspx>

Similarly, a section a section that presents facts on the growth profile of the water supply network on the CMWSSB's website, states that 100% of the city's area is covered with piped supply.<sup>16</sup> However this claim is contradicted by data in other Metrowater documents. For instance, one component of a project implemented under a centrally sponsored urban mission, the Jawaharlal Nehru National Urban Renewal Mission (JnNURM) aimed at extending the water supply network to "unserved and unsewered areas",<sup>17</sup> indicates that not all of the city is serviced by the utility and casting some doubts on the accuracy of the CMWSSB's data.

### **Incomplete data**

Data on municipal infrastructure created and maintained by the CoC do not have complete information. For instance, none of the data maintained by the CoC are geo-referenced. According to our interviews, the GIS cell, which was created through funding from the central government, mostly helps to "prepare AutoCAD files for different departments, make TIF images of Google maps, verify the EDP road list, assign IDs to the streets, print large A0 size maps."<sup>18</sup>As a result, crucial parameters that can help understand the infrastructure, such as unique identification numbers for public facilities and their locations, are missing.

In such cases, street-level bureaucrats rely on their informally-held knowledge about the city to monitor existing infrastructure and plan for new infrastructure. In one interaction with the Bridges department, we were told that the satellite images, which were screenshots from Google that the GIS department provided, were sufficient to identify locations for new bridges. They added that the actual plans were prepared by consultants and that it was incumbent on them to collect or create the necessary data.<sup>19</sup>

Similarly, we observed how complaints received for public toilets could not be tracked over time to associate complaints with specific toilets, and to investigate complaints registered for a particular toilet were problems with maintenance or if there were systemic problems such as problems with the basic toilet infrastructure. In the absence of such a tracking system we saw that complaints about individual toilets were regularly addressed through the efforts of Ward and Zone officials, but these were ad hoc and in response to each complaint. The absence of unique identification and location information had implications for planning as well. For example, when the CoC was planning for new public toilets, several locations from a proposed new locations list already had an existing toilet.

---

<sup>16</sup> Chennai Metrowater, "Water Supply System" accessed on May 26 2014, <http://www.chennaietrowater.tn.nic.in/departments/operation/developwss.htm>

<sup>17</sup> Chennai Metrowater, "Summary of Detailed Project Report for Improvements to Water Supply System in Chennai under JnNURM", accessed on May 26 2013, <http://www.chennaietrowater.tn.nic.in/departments/projects/jnnurm.htm>

<sup>18</sup> Informal discussion with NVXL, June 17, 2014

<sup>19</sup> Interview with MBCZ, December 09, 2013

Similarly, when the Storm Water Drains (SWD) department was planning for the construction of new rainwater harvesting structures to direct rainwater underground rather than to drains, they intended to use Google maps to identify locations of water stagnation.<sup>20</sup> They relied on an outdated AutoCAD map that was neither to scale nor geo-referenced to identify under-served areas. The departments worked with private consultants to plan for new infrastructure but lacked the detailed knowledge about the wards. A related problem was that the engineers did not acquire the data from the consultants and so the data gaps within CoC persisted.

### **Lack of data formats and standards**

There are few standards and formats for the way data in the CoC are organised. In fact, we found that data collection processes respond to demands for reports, rather than to ensure that agreed-upon standards and formats are maintained. Most of the data collected and used by the city engineers is organised in the format it has to be reported in to senior officials at review meetings, where monitoring of progress is largely restricted to monitoring finances. In ‘Works review meetings’ that were held at the CoC every fortnight the data presented is organised in columns titled “works completed”, “works contracted”, “to be taken up”, etc. There are no data fields that indicate whether the toilet has been assigned a sanitary inspector (the street level bureaucrat in charge of cleaning toilets) whether it is cleaned regularly, and whether it is stocked with adequate cleaning supplies. Since most of the data on toilets deals with tracking them as construction projects, there is very little focus on the quality and purpose it will serve for residents.

In interactions with staff at a CoC dispensary, we found that they were required to collect data in several registers.<sup>21</sup> The Medical Officer, for example, was required to maintain the Morbidity register, Nominal register, Slum register, School health register, Medical camp register, Injection register, Movement register and the Out-patient register. The Pharmacist maintains a Stock register and a sub-stock register, which lists items sold from the stock. The Laboratory Technician maintains the following registers:

- Sugar Register- persons who come for screening of Diabetes
- Diabetes Mellitus (DM) Register – confirmed cases of Diabetes Mellitus
- Blood Pressure Register
- Sputum Register- For Diseases such as Tuberculosis
- Malaria Register- registry of persons who are screened for Malaria
- Malaria positive Register – confirmed cases of Malaria
- Biochemistry Register
- Hyper Tension Register

Despite the presence of several formats for data collection for public health aspects, like in the Works department, the bulk of the data collected by the Health department is collected and used for reporting. In an interview with a health officer, we were told that these reports

---

<sup>20</sup> Discussions with SWD department, May 19, 2014

<sup>21</sup> Interview with UTBAWN October 31, 2013



are sent to the headquarters in different formats depending on what the information is being used for at that time, making it difficult to compare such reports over time.<sup>22</sup> In fact, our interviews with city engineers and public health officials revealed that reports are generated for specific meetings and are not saved systematically anywhere.

Since data is created primarily for reporting, it focuses more on administration of infrastructure than its quality or use. Non-standard reporting formats make data analysis a time consuming process. Most data on the CoC and CMWSSB websites, and also on several other government websites, are presented in images, or as PDF files, formats that make it inconvenient and time consuming to crunch the numbers. As a result, the data does not easily lend itself to an analysis of the extent of service provision or equity in access to civic services.

### **Selective disclosure**

Under the RTI Act, all public agencies are required to make their data public, except for information that can compromise the country's security. However, we have observed that neither the CoC nor CMWSSB followed a standard for proactive disclosure. Rather, we found that both agencies made only select kinds of data public, data inadequate for holding the water utility accountable for providing services for the poor, and in formats that did not lend themselves to ease of analysis.

For instance, one dataset<sup>23</sup> on the CMWSSB website contains some basic data on water in the city, but excludes information that could facilitate an analysis of water distribution. The data presents some facts on the storage capacity of the various lakes that service the city – the Poondi, Cholavaram, Redhills and Chembarambakam reservoirs. From these sources, water is pumped to treatment plants and distributed to the rest of the city through a network of pipes, pumps and conduits. However, for someone looking to understand why there seem to be inequalities in access to water across the city, the data that is available does not provide any insight. For example, the information provided is not enough to answer questions such as: do all neighbourhoods within the city receive equal amounts of water? How is demand for water from different areas established? Are tail end areas in one distribution system serviced by another distribution station with excess water? Answering these questions will provide insights into whether Chennai's water distribution system is skewed in favour of specific areas or populations. Unfortunately, the existing data on CMWSSB's website does not lend itself to such an analysis.

There is absolutely no data about public health in the public domain nor information on where such data can be found, an observation that has been substantiated by professionals in the private healthcare sector in Chennai.<sup>24</sup> A senior health officer said that placing data on public health in the public domain was frowned upon by the city administration and the officers were punished for sharing data with journalists and researchers.<sup>25</sup> He added that a

---

<sup>22</sup> Interview with GHJKL, October 29, 2013

<sup>23</sup> Chennai CMWSSB, "Lake Levels", accessed on May 26 2014, <http://www.chennaiCMWSSB.tn.nic.in/public/lake.htm>

<sup>24</sup> Interview with YUBFFG, November 28, 2013

<sup>25</sup> Interview with PQRST, July 19, 2013

few years prior the City Health Officer had been suspended from service because he had shared information on the number of deaths resulting from dengue with a journalist. He went on to explain that "too much data can cause an unnecessary panic" and this is why not all public health data was made public. The CoC's website provides data on the kinds of primary healthcare facilities offered in the city and the numbers and street addresses of these facilities. However, the data is outdated: It is only relevant for the erstwhile ten zones, rather than the full city, and has still not been updated more than three years after the city's expansion.<sup>26</sup> There is also information on what facilities are available at the primary healthcare centres, but these are aggregated at the city level. So if a resident wanted to know where he could avail one of the listed services - ECG, ultra sonogram, auto analyser, and binocular microscopes - they would not know which dispensary to go to.

### **Lack of metadata and licenses**

Most of the data that is available publicly is presented without metadata - contextual background that is necessary for a full understanding that is necessary to be able to use the data to understand government performance. Descriptions of the data fields that populate a dataset, information on how the data was created or how frequently it is updated do not accompany data presented on the CoC and CMWSSB websites, precluding any meaningful analysis of the information. The CoC's Public Health department requires its staff to maintain several registers to collect data - we estimate close to 20 different registers!<sup>27</sup> However, there is no metadata explaining what each of them are about or how they get linked. A senior health officer said that there were several registers and formats were collecting data but that they were difficult to understand and verify.<sup>28</sup> He had also said that linking different kinds of data would have allowed for better insights into planning and monitoring health outcomes, but the lack of metadata hindered this. A private health professional added that though the CoC gave them formats to submit their data, there were certain kinds of data that were not amenable to the formats right away. For instance, diseases such as malaria, leptospirosis, typhoid and scrub typhus have similar symptoms initially and patients need to be placed under observation for 48 - 72 hours. Doctors may either reserve their diagnosis or change their diagnosis depending on how the patient progresses and the test results. Private hospitals are still required to report these but there is no way for them to update the data and link with earlier cases or diagnosis.<sup>29</sup>

For example one of the datasets on the CMWSSB website contains information on the number of trips made by water tankers (or water lorries) to supply water to residents.<sup>30</sup> This data is important because water tankers often supply water for residents who are not connected to the piped water network. These are most often residents of slums, but

---

<sup>26</sup><http://www.chennaicorporation.gov.in/departments/health/infrastructure.htm> Website accessed on September 12, 2014

<sup>27</sup>Interview with UTBAWN October 31, 2013

<sup>28</sup>Interview with GHJKL, October 29, 2013

<sup>29</sup>Interview with YUBFFG, November 28, 2013

<sup>30</sup>Chennai CMWSSB "Status of City Mobile Water Supply (Lorry)", accessed on May 26 2013, <http://www.chennaiCMWSSB.tn.nic.in/public/lake.htm>

sometimes also other residents whose water supply is temporarily interrupted. This dataset is downloadable in excel format and updated every day. It records the number of water tankers that supply water to the city every day, the capacity of each lorry, the number of “payment trips”, and the number of “vehicles allotted”. However, this dataset is not accompanied by any explanation of the data in its neatly formatted rows and columns. One can only speculate that the data on “payment trips” refers to the number of water tankers that deliver water on demand to marriage halls, government offices, schools, colleges and commercial establishments, but this is not clear, and repeated enquiries to the water authority did not clarify the meaning of this category.

Additionally, none of the datasets that Transparent Chennai obtained from the CoC or CMWSSB over the years 2010 – 2014 had any license. Often when information was obtained through requests to government officers that Transparent Chennai researchers knew, they were told that the data was given in confidence, and that it could not be published or given to the media. When we sought permission to interview CMWSSB engineers and requested particular pieces of data, there was a caveat that we were strictly prohibited from sharing or publishing the data, and publishing any derivative had to also first be approved by a senior official at the agency.

### **Negligible sharing of data**

The poor quality of data is also affected by the near-absence of sharing of data between the city government and the water agency, and between public and private healthcare professionals. Both the CoC and CMWSSB maintain datasets on toilets, slums and tax assessments. However, they maintain these separately and the datasets are not comparable. The numbers and locations of toilets and slums in these datasets differ, as do details of the tax assessments. Moreover, some of the data from the two entities is contradictory, casting doubts on the quality of data from both.

For instance, CMWSSB has a separate list of toilets in the city. TC filed an RTI petition on 28 October 2013 to access CMWSSB’s data on public toilets. TC has not received responses from all the areas, but from those received it is clear that number of toilets in the zones reported by Chennai CMWSSB and CoC do not always match.

<b>Zone Number</b>	<b>Corporation of Chennai</b>	<b>Chennai CMWSSB</b>
Zone 6	80	80
Zone 7	22	24
Zone 9	137	115
Zone 10	64	51
Zone 14	17	26
Zone 15	22	16

**Table 1: Numbers of toilets according to CoC and CMWSSB**

The Corporation of Chennai assigns a number to all private residential property to calculate and recover property tax. The CMWSSB Board calls this number the “Chennai Municipal

Corporation” or CMC number and requires residents applying for water connections to specify their CMC number on their application forms.<sup>31</sup> The Board uses this number to identify eligible applicants, estimate demand for water connections, and calculate its revenues. Even though the CMC number has immense significance for public access to water – it determines an applicant’s eligibility for a connection – there seems to be no consistent system for this data to be updated and shared between the agencies. One official revealed that the CoC sends a list of CMC numbers to CMWSSB every month, but that “it becomes difficult to manage this data.”<sup>32</sup> He said that many records get duplicated and some old records get deleted in the database when updated.<sup>33</sup> Another official in charge of a different area said that the data “is received from the Corporation in a CD only once in six months.”<sup>34</sup> What is evident is that there is no established way in which data is shared between the two entities.

The CoC collects so much data from private healthcare providers but there is no information on how it is being used. There are very little joint activities or few that leverage private service providers’ strengths. Patients are required to continue treatment at home and come back to the hospital for follow-ups but often do not, and in the absence of a link between public and private databases, it is difficult to track the patient. The public health ramifications for the city are enormous but the city’s ability to pre-empt epidemic outbreaks is weak.<sup>35</sup>

### **No integrated repositories of data**

There is no single place for all the data for the CoC is stored. The Enterprise Resource Planning (ERP) software used by the CoC has thirteen ‘Works’ modules, all of which are anchored in the financial transactions of the works undertaken. The software allows officials to create and monitor contracts for infrastructure construction and data exists only if there was a financial transaction, but related data, such as the quality of the asset created or the quality of work by contractors, is held informally by different officials and is not linked to the work listed in the ERP.

In interviews with city engineers, we were told that the system is complicated to use and inadequate for the kinds of data they are expected to collect. Moreover, despite having tools for monitoring the progress of projects before issuing payments to contractors, such provisions are easily bypassed. One ward engineer told us that after an estimate for the creation of an asset is approved, the system requires that the engineers track the progress of the contractor by setting reasonable “milestones”. These milestones are the percentage of the project completed by a certain date. Calculating the percentage of a project completed on the ground is difficult, and the engineer revealed that he often updated this

---

<sup>31</sup> “Chennai CMWSSB Application for Water/Sewer Connection”, [http://www.chennaiCMWSSB.tn.nic.in/pdf/wscon\\_eng.pdf](http://www.chennaiCMWSSB.tn.nic.in/pdf/wscon_eng.pdf), accessed May 19 2014

<sup>32</sup> Interview with BERIUP, November 12 2013

<sup>33</sup> Interview with BERIUP, November 12 2013

<sup>34</sup> Interview with AAPETS, November 5 2013

<sup>35</sup> Interview with YUBFFG, November 28, 2013

on the system without actually visiting the project site.<sup>36</sup> Another official lamented that the system never captured the exigencies of the moment: the rising price of cement, unexpected rain and absentee contractors – factors that actually explain delays and increases in project costs. “The head office only wants to know whether the project that was taken up is over,” he said.<sup>37</sup>

Although the Works system focuses on infrastructure, it fails to record complaints about infrastructure. To do this, the CoC **maintains** the Public Grievance Redressal (PGR) system. The **PGR** system collects complaints in a variety of formats – by SMS, email, on the phone and by letter. The PGR system automatically assigns a number to each complaint, forwards it to the concerned engineer, and provides the complainant with status updates.<sup>38</sup> The system has an element of interactivity to it – citizens can reopen a complaint several times, and send messages to the official in charge of responding to it. There are other complaint portals that operate in the city as well – the office of the Chief Minister of Tamil Nadu recently launched a website and SMS facility for residents of the state to submit their grievances and petitions. Complaints about issues that fall within the jurisdiction of the CoC are forwarded to the CoC’s redressal system.

Since the two systems are not integrated, the CoC does not capture information on whether the same toilet is being complained about repeatedly, making it impossible for both residents and the government to hold an engineer accountable for bad maintenance. Further, none of this data is not open to the public, so it is not possible to monitor whether the complaints are in the process of being resolved.

### 3. Discussion

**Does the evidence collected suggest that government data quality is indeed questionable? If so, what are the ways in which this data is deficient, particularly with regards to the poor?**

Our research findings reveal that data maintained by the CoC is inaccurate and incomplete to a large extent. Both CoC and CMWSSB also have published datasets on their websites that demonstrate poor data standards and formats, and lack metadata and licenses, making it difficult for citizens and researchers to access and analyse data. The quality of disclosure by both agencies shows that they are extremely selective in the kinds of data that they proactively release.

These deficiencies are exemplified in the CMWSSB data on the number of water and sewer connections. While the agency claims that 100% of the city has been covered by piped connections, in JnNURM-related documents there is a mention of improving access to unserved areas. A similar inconsistency can be observed when comparing the number of water and sewer connections under the Economically Weaker Section scheme that aims to

---

<sup>36</sup> Interview with UFNDGE, August 8 2013

<sup>37</sup> Interview with UFNDGE, August 8 2013

<sup>38</sup> Demonstration of the PGR system, nd.

improve access to low income groups. Under this scheme, the agency states that a water connection can only be obtained if the applicant has a sewer connection but the difference in numbers contradicts this position.

Similarly, the CoC did not have an accurate understanding of access to water and sanitation, and even populations in slums in the city. It also does not have information on street vendors, informal transport service providers, illegal sewer and electricity connections. These gaps in data result in the CoC providing public services, such as public toilets, but these are often inadequate and inappropriately located. We can observe toilets clustered where there was land available rather than where there was a lack of private sanitation facilities, or hubs that attract large transient populations such as informal markets and informal transit locations.

### **Do these deficiencies result from common failures in data collection and storage processes?**

Data on demand - this is what the head of the Works department of the Corporation of Chennai had to say when we asked him about what data exists with the civic body.<sup>39</sup> He had recently taken office and found himself handicapped with the complete absence of data and the underlying lack of data collection formats. Rather than follow existing formats for data collection, city engineers tended to collect data depending on reporting was required for specific meetings. They sent these reports in non-standard formats and did not save it for posterity. Apparently neither did the officials at the headquarters save the data because ward and zone officers confided that they were often asked to send the same data multiple times.

The deficiencies in data within the CoC are compounded by not having easy access to data from other public agencies. We faced a lot of resistance from various government agencies when we approached them for data and were given several reasons to justify their denial of access. These included reasons such as the data are too big to share; it is sensitive, people will misunderstand it, and another department has it (and if you get it, would you please give it to us?).<sup>40</sup> What makes it further difficult to plan for municipal infrastructure is the fact that several government agencies are required to coordinate with each other to ensure that the facilities are functional. However, this necessarily means that the agencies need to share data about the infrastructure and services.

Public toilets, once constructed by the Buildings Department of the CoC, have to be connected to the city's sewer network, a network operated by another government agency, the Chennai Metropolitan Water Supply and Sewerage Board (CMWSSB). Electricity connections to toilets are provided by the Tamil Nadu Electricity Board (TNEB). Each agency collects and stores different data on toilets, and while the CoC is responsible for planning for some public services in the city, it rarely has a complete picture of the

---

<sup>39</sup> Interview with Z1705, August 1, 2013

<sup>40</sup> Data bingo exercise by the Transparent Chennai team

information required to plan for how many public toilets are required and where they should be located, because data is not shared between the agencies.

**Does this poor quality of data actually impact the quality of services provided to residents, particularly the urban poor?**

The city government and the parastatal agency are both responsible for increasing and improving access to municipal services and infrastructure, with a focus on the urban poor being a desirable feature of public service provision. However, because the government lacks the data necessary to plan and monitor public service delivery clear impacts on the ability of the city government to provide such services, particularly to the poor. The locations of the proposed new toilets, for instance, did not account for where existing public toilets were located. From the virtual mapping and physical survey exercise we conducted, we found that the JEs who had proposed the locations had not considered the presence of existing toilets before proposing new toilets! Had the tender for the new toilets been successful, the project would have resulted in a crowding of toilets around certain areas and other areas remaining underserved. Leaving the location of the toilet to the company's discretion would have also resulted in a situation where they would locate toilets in areas where advertisements were lucrative, rather than where toilets are most needed.

The lack of coordination and sharing of the data between agencies also impacts how residents have access to public sanitation. For instance, engineers at the CoC are responsible for fixing blockages in the length of pipe between the toilet and the main sewer.<sup>41</sup> It is Chennai CMWSSB's responsibility to ensure that its sewer mains function properly.<sup>42</sup> If there is a problem with a public toilet's sewer connection, the Junior Engineer at the CoC has to contact an engineer at CMWSSB to fix the problem.<sup>43</sup> On the one hand, officials from the CoC allege that CMWSSB often takes more than their stipulated time of three days to fix a block in the sewer,<sup>44</sup> and the toilet is unusable during this time. On the other hand, Chennai CMWSSB officials feel that the CoC should be more equipped to handle toilet maintenance. One CMWSSB official said that while Chennai CMWSSB is responsible only for clearing obstructions in the sewer mains, the agency ends up executing plumbing work in toilets as well, because the officials at the CoC are not equipped with basic skills. One official admitted that the CoC "must clear minor blockages themselves."<sup>45</sup> Another said, "CMWSSB cannot act like the problem solver for all government agencies."<sup>46</sup>

The coordination between the local government and Chennai CMWSSB, affects not only access to water and sanitation, but also how the city is planned. The CoC and CMWSSB have to coordinate closely on "road-cutting", the act of digging roads to maintain, install or remove underground infrastructure like pipes, sewers and cables. In order to make a

---

<sup>41</sup> Interview with DHDJCK, August 21, 2013

<sup>42</sup> Interview with EQRIUY, November 21 2013

<sup>43</sup> Interview with AEVIDF and HHFAAC, January 30 2013

<sup>44</sup> Interview with UFNDGE, August 8, 2013

<sup>45</sup> Interview with HEAWEQ, August 8 2013

<sup>46</sup> Interview with FRESOA, November 8 2013

road-cut, government agencies have to submit an application to the CoC, as the local government owns most roads in the city. Even though coordination between the CMWSSB and CoC happens at formal and constitutionally mandated Ward Committee meetings held at the zone office every month, there are no standard data formats and lapses in data sharing. As a result, the CoC is not always apprised of when and where the CMWSSB digs up roads. One official at the CoC said that “CMWSSB are the kings” and admitted that they dig up roads at night, implying that they did not bother to seek permission. One official at CMWSSB admitted to problems with coordination and information sharing. “Periodically the Corporation informs us about their plans, he said. But sometimes they ‘do not plan’. “You know when an important official is passing through the road, it is laid overnight”, these things no one can plan for.”

The deficiencies in the quality of data, storage and sharing have similar impacts on nearly all other services that are primarily accessed by low income groups - public transport, footpaths, government schools and hospitals. Most of these infrastructure and services are provided by different public agencies in the city and require a high degree of coordination. However, the absence of data formats and templates, poor data and negligible sharing reinforces the fragmentation in governance. But most importantly, the deficiencies of data have important ramifications for accountability - both internal accountability and the ability of citizens to hold governments to account.<sup>47</sup>

### **Are there ways in which this data can be easily corrected or improved?**

In the example of public toilets provided above, it was only when we worked with the CoC to create data collection methodologies that it was possible to create better data and to involve the public in decision making. What started as an examination of the state of data on public sanitation and the ways in which they are used for decision making, Transparent Chennai's engagement with the city government changed the parameters of data collected, the methodology for collecting data and incorporating it in the tender document. More importantly, we were able to shift the objective of the provision of public toilets to the need for sanitation facilities rather than a political announcement or business interests of concessionaires.

In the process of our research engagement, the CoC signed an agreement to create a Chennai Data Portal project team that would work to improve data management practices on a range of issues. As part of this collaboration, we worked with different engineering departments to articulate desirable outcomes of providing public infrastructure and services, and indicators by which their performance can be measured. This entire intervention resulted in the institutionalising of the creation and use of data for planning and monitoring of public toilets, something that is being replicated for all other infrastructure provided by the city government: bus shelters, footpaths, parks, playgrounds, roads and stadia, to name a few. We followed this with data collection exercises, including building violations, trees, and pedestrian-vehicular conflict.

---

<sup>47</sup> Interview with Z1705, August 1, 2013



One significant example is that of creating a micro plan when heavy rains inundated parts of the city in October 2014. CoC officials felt compelled to acquire accurate information about the location and reasons for the water stagnation in order to create a plan with interventions that addressed each location. Following an exercise where city engineers and other staff mapped the locations where water stagnation was observed, we worked with each ward engineer to document the plausible causes and the potential immediate and longer term measures that could be used to prepare a micro plan. The plan also included a cost and time estimate associated with each water stagnation location. This exercise revealed that Ward engineers were aware of the problems, which included incorrect road cambers, blocked storm water drain inlets, misalignment between storm water drains and other channels, illegal connections between storm water drains and sewerage lines. However, they were not included in the planning exercises of these municipal infrastructures even though they were ultimately responsible for their maintenance.

## 4. Conclusion

In contrast with Transparent Chennai's civic-engagement approach is that of the 'smart cities', which is made up of corporate-led projects which focus largely on providing world-class infrastructure. The approach is almost entirely based on the collection of data using cameras, sensors and other technology, and the use of technology to support decision making. Unfortunately, the projects focus on infrastructure and services, such as transport, water and electricity, which are more amenable to the use of meters and GPS devices for collecting data. The risk of this is that it excludes all users of these services who do not have piped water and sewerage connections where meters can be installed or services providers, such as informal public transit services, who do not have GPS devices. In general, it belies the realities of Chennai, and several other cities, where informality is pervasive and often the way that citizens access and provide services.

In an attempt to democratise data and promote transparency and accountability, India has enacted the Right to Information Act, 2005, implemented the National Data Sharing and Accessibility Policy (NDSAP) and also set up an open data portal that publishes government data.<sup>48</sup> However, as this report highlights there are several factors that need to be considered before concluding that open data could unequivocally result in the benefits associated with it. Given the state of data in India, there are several problems with its quality and opening up inaccurate data to the public could have several undesirable consequences. While on the one hand it could encourage people to challenge official data, on the other, it could result in a reinforcement of its inaccuracies and inconsistencies, both in collection as well as use of inaccurate data. Along with improving the quality of data, it is important to work with their users, government and residents, to improve the way data is collected, organised, stored and shared. In this, the role of intermediaries can be to help bring in data standards, formats, tools and methodologies.

---

<sup>48</sup> Department of Science and Technology, "Introduction to the National Data Sharing and Accessibility Policy (NDSAP) ", accessed May 29 2014, <http://www.dst.gov.in/nsdi.html>

