

Data, Urbanisation and the City

ANANT MARINGANTI, PARTHA MUKHOPADHYAY

By using the enormous processing capacity of computing that is now available, we can, it is claimed, improve how cities are governed—make them smart! This review attempts to illuminate how data reveals relationships between citizens and the state and thus facilitates an informed debate on whether data can be deployed to build a more inclusive and constructive relationship between citizens and their government. As urbanisation deepens, we see struggles around who gets to decide what is to be governed and how the data is to be collected and deployed and what technologies and skills are to be deployed for implementation. The papers in this collection can be viewed in three groups, respectively, dealing with three issues: data collection processes, intra-urban spatial inequities and use of new sensing technologies.

The ubiquitousness of computing power now makes it possible to collect and process very large amounts of data in a relatively short time. Using this processing capacity can, it is claimed, improve how cities are governed—make them smart! Over the last year, the idea of smart cities has gained currency in research and policy debates on Indian cities. The launch of the smart cities scheme by the Government of India has given more concreteness and urgency to the debate on what kinds of socio-spatial consequences such schemes can engender. This edition of the Review of Urban Affairs takes a step back from the immediate debate on smart cities to look at the use of data in and of cities; to both broaden the field of inquiry and delimit the scope of debates on India's urbanisation.

Conjuncture of Three Developments

The current moment in India's urbanisation can be seen as a conjuncture of three interrelated developments.

First, there is increased availability of national- and city-level data sets with some degree of spatial information. This data is largely generated by national, state and city government agencies. Their availability to non-state actors is fragmented, regulated by legislation, by political and policy expediency and sheer bureaucratic inertia. Yet, since the 2011 Census, more such data has become available. It has become possible to gain insights into the spatialities of Indian cities and urbanisation processes in new ways by overlaying data sets from different origins. Increasing confidence among researchers on the potential of these data sets can lead to more vocal demands for release of data sets created by government agencies.

Second, new types of data are being generated through the use of passive sensing technologies, that is, generated and or accessed from technologies that are already in use. For example, the use of smartphones generates information about location and search behaviour. Digital grievance redress systems can generate enormous volumes of data on individual and group behaviour. These data sets are being harnessed by civic groups, enabling a new generation of politics around accountability.

Third, there is increased access and sometimes control of these data sets by information technology (IT) professionals, planners, researchers and commercial entities, that is, market and civil society actors as distinct from public agencies. A number of questions arise as a result. How reliable is the data? What is the data used for? How can data be used to ask relevant questions about the opportunities and challenges of Indian urbanisation? In sum, how can we be smart about the city and the process of urbanisation?

Anant Maringanti (amaringanti@gmail.com) is at the Hyderabad Urban Lab; Partha Mukhopadhyay (partha@cprindia.org) is at the Centre for Policy Research, New Delhi.

Two data-related issues will be conspicuously absent in this review—primarily because both of them require dedicated attention by themselves.

First, the use of spatial data at the regional or larger scales to investigate settlement patterns, their typology and distribution across space. Recently, the Registrar General of the Census of India released data on household characteristics at the settlement (both village and town) level across many dimensions. However, neither maps nor more limited geo-references of these settlements have been officially released, though it is available in private data sets. This limits analysis of the spatial character of development across these dimensions, which is related to the current debates on the rural/urban/“rurban” questions. These data-related inadequacies require open discussion in their own right.

The second issue, in our view, which demands dedicated attention in the immediate to medium-term future in the Indian context as new technologies penetrate Indian economy relates to the way in which data technologies are being used to build connections among citizens in a relatively anonymous and yet interactive manner. Two uses, in-city taxi substitutes (e.g., Uber) and short-term rentals, that is, hotel substitutes (e.g., Airbnb), currently dominate the so-called “share economy,” which should perhaps be more accurately called the “pay to share economy.” They enable citizens to transact among themselves, with only digital intermediaries. Thus new providers enter and exit the market on a dynamic basis. Depending on behaviour, this can reduce or increase volatility of services in the market. It also redefines the relationships of employment and labour, and challenges existing modes of regulating these services. For example, the rise of Uber in the provision of taxi-like services has significantly reduced the number of taxi rides in cities like San Francisco and led to legal challenges in New York.¹ It is interesting that in San Francisco, the mayor’s office tried to solve the problem of underutilised taxis but this attempt at becoming smart was overwhelmed by a radical change in the transport market engendered by the arrival of Uber, which has led to a 65% drop in taxi trips.

Similarly, the hotel and housing rental market now has to face competition from apps like Airbnb, which have brought in new supply into a rental market that did not exist before. However, this kind of technology also makes a different kind of business possible, where people invest in houses, solely for the purpose of putting them up for rent on such sites. In the context of the United States, a number of questions have arisen. Is there a difference between hosted (where the owner is staying in the house) and un-hosted stays, which are likely to be more of a commercial nature? Should there be limits on the number of days a house is used for such purposes? Do such persons become landlords, with the responsibility that status brings? Do people staying there have the protection of tenants? At a broader level, is the public housing supply affected if the supply of low-income rentals fell as a result of people withdrawing from that market to use their houses for short-term rentals?

In addition to these, there are tax-related issues—both with the taxability of income and the tax implications of using a property for commercial purposes, as in Airbnb and the regulatory

regime for transport services. Since there is a lot of local variation of laws in the us, many cities, counties, and other municipalities have legal restrictions on short-term home rentals, while others are less constraining. In the case of car services, there is the question of loss of medallion (a licence to pick up customers on the street) sales for cities, as the value of a medallion drops due to use of apps like Uber. There are also issues related to driver safety and insurance, as in India. In part, there may be a lack of compliance due to lack of knowledge of various tax laws by first time entrepreneurs and in part due to deliberate tax avoidance.

All of these issues have been raised in various litigations and regulatory decisions in the us in cities such as San Francisco, New York, etc. The “share economy” is currently in a regulatory grey zone and may even foster the growth of an underground, informal economy in countries such as the us, but a number of new rules are being written to deal with changes brought about by technology. Even as this happens, even without the regulatory clarity, Uber is reportedly valued at \$40 billion, that is, more than Infosys and Airbnb at \$20 billion, which is more than WIPRO.²

One key feature of these applications, which could have significant implications for more public services, is the ease with which personal information is shared with third parties, in this case with corporate firms. Does this ease of sharing of information with the government voluntarily—since one has already shared this with private providers—thereby make a panopticon state more feasible? In any event, with new legislation like the Patriot Act in the us and the recently legislated laws in France, governments would have the right to procure these records from these service providers. By participating in the share economy, have we already started easing panopticism gently into our lives? This is especially concerning for India, which has yet to enact any form of privacy legislation to prevent misuse of data.³

Goal of Collection

The central goal of the collection of papers in this review is to illuminate how data reveals relationships between citizens and the state and thus help us create an informed debate on whether data can be deployed to build a more inclusive and constructive relationship between citizens and their government.⁴ As urbanisation deepens, we see struggles around who gets to decide what is to be governed and how the data is to be collected and deployed and what technologies and skills are to be deployed in implementation. The papers in this collection can be viewed in three groups, respectively, dealing with three issues: data collection processes, intra-urban spatial inequities and use of new sensing technologies.

Urban researchers in India are familiar with the fact that one of the most common, albeit non-transparent, uses of data is in slum relocation and rehabilitation, where data on location,

The Review of Urban Affairs’ external editorial advisory group consists of Partha Mukhopadhyay, Anant Maringanti, Vinay Gidwani, Karen Coelho, Amita Baviskar and Ashima Sood. We gratefully acknowledge their help in developing the thematics of the RUA, identifying potential contributors and reviewing contributions.

number of households, length of stay, etc, are collected. As inequity is reproduced and aggravated not merely through policies but also through administrative processes, errors and contestations of such data have significant human consequences. In particular, a basic issue is about the nature of the data gathering process, its collection and verification—and how does the data inform public action, and if so, how does it do so? What happens when action is taken by those who are not fully familiar with the contextual nature of the data? Sheikh and Banda (p 73) in this review look at both the various processes that generate data, their legal sanction as also the actual data on *jhuggi jhopdi* clusters (JJC) made available by a more transparently functioning agency, Delhi Urban Shelter Improvement Board (DUSIB) over time, and ask questions about its quality and consistency. They reveal how the move to a more data-driven decision process can both exacerbate a number of older issues relating to uses of data in our cities—the silences and invisibilisation processes that arise from them and yet open up new ways to make visible those very processes with greater transparency.

Indian urban planning has historically been largely confined to the generation of land-use maps and zonal regulations and has paid limited attention to services. This practice makes service invisible and allows a free pass to both ill-planned resettlement colonies for the poor which, like Savda–Ghewra in Delhi, are developed without service provision, as also provisions for planned sub-cities like Dwarka, which was settled without minimally adequate water supply.⁵ The consequential urban inequity which has built up over decades, not just in Delhi but in almost all cities, through layer upon layer of infrastructural and other investments both by the state and by non-state actors, including individuals, has completely gone undetected because of lack of granularity and spatial referencing of data. Three papers in this review use the newly released ward-level data on household characteristics to begin to make some of this visible. Balakrishnan and Anand (p 63) constructs a methodology for generating neighbourhood typologies (at the ward level) for Bengaluru. Bhan and Jana (p 49) use this data to ask questions about the spatial variation in wealth and housing across the city, with a focus on Bengaluru and Delhi. Similarly, Sidhwani (p 55) looks at the spatial distribution of caste and its relationship to

specific services. The insights that emerge reveal a much more layered and complex Indian city than what was visible before such data was available. The authors also question the manner in which new data practices can lead to opportunities as well as challenges.

'Smart Cities'

The ability to process and use mobile network big data (MNBD) and other large data generated by public agencies, related to management information systems (MIS) for services, like water supply systems and electricity grids, and incident reporting, like police reports and citizen grievances, has increased exponentially over the last few years, and it is this use of data in cities, pushed hard by information technology majors, that urban managers now associate with smart cities. Samarajiva et al (p 42) look at how MNBD can reveal some aspects about large cities in a near real-time manner, especially about dynamic aspects like transportation, that was not possible with the traditional survey and census techniques. They illustrate this using Colombo and speculate on what kind of institutional form would be necessary to really make this available for the public good.

Ultimately, the debate around smart cities is about cities making smart choices. This debate then must focus on where cities are deploying their efforts. For example, in waste management—we could ask whether the bin is overflowing (citizen centric)—or whether the contractor's trucks are going to the dumpsite (service-centric). It must also focus on what is prioritised—tracking how much untreated sewage is flowing out or failures of electricity transformers or when the next bus is arriving or where is the next empty parking space? It must make political and social choices. How much use is made of public networks and crowd-sourcing and how much is a technical solution that is a multi-billion information technology contract? What is the mix of software and systems interventions vis-à-vis hard infrastructure—roads that repair themselves, buildings that suck pollution, concrete that heals itself, etc? These questions will ultimately force us to rework and reformulate well-researched issues on technology and society, and relate to technologies, life-cycles, the tensions between domestic and foreign sources, and expenditure allocations.

NOTES

- 1 So far, Uber, a major mobile app, has successfully argued in some courts that its service is different from a medallion cab, which can be physically hailed on the street—though it is arguably providing a digital hailing service. Indeed, in New York city, there has been a major legal battle with apps such as Uber and Hailo (a similar app for hailing taxis, which initially began in London) being challenged in court, then first declared illegal and subsequently reinstated on appeal, largely on the grounds that a number of areas were poorly served by street-hailed yellow taxis. Currently the Taxi and Limousine Commission of New York is piloting a number of taxi e-hailing apps, which try to level the playing field between digital hailing and physical hailing.
- 2 Not just taxis and hotels, there are now a number of such apps for eating out, renting cars and houses, renting tools, etc. As more people seek services from individuals instead of establishments, the nature of the urban

economy could change radically. Added to this is the possibility of disruptive decentralised technologies like 3-D printing. The manner in which the city can change as a result is beyond the scope of this paper, but worth investigating. In India, these app features and demands are limited to a relatively small population, even though the number of internet users through mobile phones may exceed the numbers who use it through a PC. Interaction between individual service providers and consumers is still individualised, but with the assistance of mobile technology. Will this morph into a "share economy" or will it stay, as with much of the rest of the Indian economy, in an individual, informal mode?

- 3 One example of privacy restriction in New York City is that the city cannot ask the utility provider for the gas and electricity consumption of each building. So, it has mandated compulsory reporting for buildings above a certain size, where buildings are required to report consumption. However, this report cannot be cross-verified with the utility records.

- 4 Here, we ignore the issue of non-citizens. Many cities will have people living in them who are non-citizens, typically foreign nationals who do not elect the government. However, this also finesses the issue of large number of within-country migrants who may not be able (or even willing) to transfer their voting rights from their place of origin to the city.
- 5 In part, this is because the development of the sub-city was with one agency, the Delhi Development Authority (DDA), reporting to the union government and the service provision of water with another, the Delhi Jal Board (DJB), reporting to the Government of the National Capital Territory of Delhi (GNCTD). The original plan was to supply Dwarka with savings in canal water transmission from Haryana, consequent upon lining the canal. Until recently, Haryana had reduced the release into the canal, thereby appropriating the savings for itself rather than GNCTD, which paid for lining the canal.