Development Informatics

Working Paper Series

The Development Informatics working paper series discusses the broad issues surrounding digital data, information, knowledge, information systems, and information and communication technologies in the process of socio-economic development

Paper No. 77

Urban Slums in a Datafying Milieu: *Challenges for Data-Driven Research Practice*

BIJAL BRAHMBHATT, SIRAZ HIRANI, NEHA LAL & BHUMIKA CHAUHAN

2019

Published in collaboration with, and with the financial support of, the University of Manchester's Sustainable Consumption Institute

Published by: ***Centre for Development Informatics***
Global Development Institute, SEED
University of Manchester, Arthur Lewis Building, Manchester, M13 9PL, UK
Email: cdi@manchester.ac.uk Web: http://www.cdi.manchester.ac.uk

View/Download from:
http://www.gdi.manchester.ac.uk/research/publications/di/

Table of Contents

Urban Slums in a Datafying Milieu: *Challenges for Data-Driven Research Practice*

Bijal Brahmbhatt
Mahila Housing Trust

Siraz Hirani
Mahila Housing Trust

Neha Lal
Mahila Housing Trust

&

Bhumika Chauhan
Mahila Housing Trust

2019

Abstract

With the ongoing trend of urban datafication and growing use of data/evidence to shape developmental initiatives by state as well as non-state actors, this exploratory case study engages with the complex and often contested domains of data use. This study uses on-the-ground experience of working with informal settlements in Indian cities to examine how information value chains work in practice and the contours of their power to intervene in building an agenda of social justice into governance regimes. Using illustrative examples from ongoing action-oriented projects of Mahila Housing Trust in India such as the Energy Audit Project, Slum Mapping Exercise and women-led climate resilience building under the Global Resilience Partnership, it raises questions about challenges of making effective linkages between data, knowledge and action in and for slum communities in the global South by focussing on two issues.

First, it reveals dilemmas of achieving data accuracy when working with slum communities in developing cities where populations are dynamically changing, and where digitisation and use of ICT has limited operational currency. The second issue focuses on data ownership. It foregrounds the need for complementary inputs and the heavy requirement for support systems in informal settlements in order to translate data-driven knowledge into actionable forms. Absence of these will blunt the edge of data-driven community participation in local politics. Through these intersecting streams, the study attempts to address how entanglements between southern urbanism, datafication, governance and social justice diversify the discourse on data justice. It highlights existing hurdles and structural hierarchies within a data-heavy developmental register emergent across multiple cities in the global South where data-driven governmental regimes interact with convoluted urban forms and realities.

A. Introduction

Of late, governance regimes in the global South have been increasingly enamoured by the reach of data, digitisation and documentation in order for the state to see, regulate and govern its citizens (Scott 1999). In *Benign Dataveillance - the New Kind of Democracy?*, Payal Arora (2019) illustrates some of these rising impulses through the ongoing biometrically-driven and database-led Unique Identification Number (UID) project in India for citizen documentation; and the Social Credits system in China that crosslinks citizenry's online and offline activity to rewards or punitive action on the basis of their credit scores. While such reconfigurations within governance as a way of producing level playing fields and pushing the agenda of equity remain questionable, what continues to intrigue is the way in which the process of datafication is being made an indispensable precursor to the discourse on the future of governance.

As one of the biggest digital projects in India if not the world, UID, popularly called Aadhaar[1] is hailed as a technological overhaul of welfare scheme targeting and direct benefit transfers. Scholars argue that "Aadhaar may not only enable efficient design, delivery, monitoring and evaluation of services…[..]..but may also offer the possibility of using modern data analytics and machine learning techniques for finding large-scale correlations in user data. This, in turn, may facilitate an improved design of social policy strategies, including targeting, and early detection and warning systems for anomalies" (Banerjee 2017). Yet at the same time, the scheme has courted both debate and protests on issues of privacy, security and exclusionary biases.

Likewise in specific relation to urban governance, the rally towards data-driven decision making is being pitched as a means to reform policy and better deliver on urban development goals. The Smart Cities Mission launched by India's central government in 2015 is one such initiative towards using big data for urban governance and administration, with ambitions spanning from rationalising traffic routes through GIS to predictive profiling for surveillance. Given that the world of big data is known to cast such long shadows, why does it continue to have such resonance with decision-making bodies? What hopes are we pinning on data for fairness of outcomes and the future of equity?

Anticipating Risk? Thinking Data?

We face unprecedented rates of urbanisation in the world, with cities across Africa, Central and Latin America, and most of Asia carrying the lion's share of this burden. This brings an urgent need for mitigation against risk susceptibility and for social equity, without which the sustainability of these cities will fall into question. India, China and Nigeria will cumulatively account for 35% of the projected growth in the world's urban population between 2018 and 2050. In absolute numbers, 2050 projections estimate that India will add another 416 million urban dwellers, while China will contribute 255 million and Nigeria 189 million (UN 2018). In addition to a rising urban population, demands for energy and water will also be

[1] Aadhaar is based on the etymological play on the word 'basis' or 'foundation' with Sanskrit origins. For citizens in India, it is a biometrically-tagged 12-digit number unique to each individual that is being designed as a way to identify and document persons.

peaking and that too in much shorter time frames. By 2030, global demand for energy and water will see an expected growth of 40 and 50 per cent respectively (UN 2016).

While current estimates point out that 1 in 4 will live in a slum by 2030, this scenario is bleaker for developing countries where inadequacy of shelter will be a reality for every third urban resident. Hence the urban poor residing in informal settlements and slums in cities of the global South are highly vulnerable to resource scarcity and risk. Whether it may be incremental changes and stresses like heat stress, pollution, poverty, induced weather extremities or natural disasters, what makes the urban poor in slums particularly susceptible to risks is "lack of secure land tenure, lack of access to basic infrastructure and services, and overcrowding" (World Bank 2011:45).

For those on the veritable margins, data is being heralded as the tool that can give them visibility in discourses where they may have remained so far unaccounted. In the case of India with a population of 65 million slum residents, the potential for producing actionable data to reduce poverty and effect change is immense. After over six decades of tepid responses of centrally-funded welfare schemes and poverty reduction measures, countries like India are stepping into the domain of data with the hope and intent of catalysing new modalities of urban governance and of addressing the scarcity and risk that these vulnerable populations will increasingly face. While the promise of ICTs as a fast mover was first envisioned some decades ago, the current push for data-driven governance and 'marking' citizens has been galvanised by the promise of digitising platforms, services and records/archives. For 'invisible' slums, whose insecure tenure and non-networked nature may often keep them off official maps and city plans, this could provide the possibility of a crucial break.

We already know, though, that achieving the potential of urban data will not be all plain sailing. Ananya Roy (2009), for example, notes exception, informality and invisibility as not gaps within but modalities of planning itself. In cities of global South, these often come to be used as purposive instruments to manage the failure of planned developmentalism. To demand visibility is thus a political negotiation against exclusionary discourse of planning and policy. While invisibility may provide temporary tenure and allow life to go on under the nose of the state; to make oneself apparent, to demand or negotiate for visibility is a political move against being enumerated on terms in which one may have no say (Warner 2013). Thus slums confront the city and often remind it of many cities that live within it.

Creating and collating data about slums and grounding the origins of this data with the community can push the envelope by providing visibility and adding further granularity to what is known about these areas; an essential foundation to proofing them against a future of risk. Yet government and municipal authorities have been faced with the temptation to enumerate and mark residents of informal settlements in limited ways, and the challenge of going further than this. Linnet Taylor (2017) argues that in low-income environments accuracy and granularity of data capture remains a practical difficulty for government authorities and urban local bodies. This is very much a reality for Indian cities which have high density slums that are constantly under the process of incremental building and have a high rate of attrition (Ghertner 2016).

So, while the promise of data-driven decision-making is real, what does this mean in everyday practice? First, as just noted, there are challenges in the very process of collecting and analysing data, such as the complexity of capturing the details of dense slums in Indian cities and the political implications of this process. But more importantly, even if such data can be gathered, there are real-time dilemmas and impediments in making use of it for evidence-based civic participation. Thus, analysing the issues that plague this practice especially for communities residing in informal settlements in the global South not only has potential to improve operational efficiency of processes but also to better deliver on goals of social equity and justice.

Mahila Housing Trust (MHT) is an Ahmedabad-based NGO that has been working with slum communities for over two decades in areas of habitat upgradation, improving livelihoods and productivity, and building climate resilience. Over the course of the last 25 years, MHT has attempted to create both broad-range and particularised data to understand the demographics and deliver on the needs of communities and households in informal settlements across eight states in India. With a deeply embedded community-based and community-led model, MHT's on-ground practice attempts to realise this process of moving from data to knowledge and subsequently to action not only 'in' but 'with' slum communities.

In partnering with communities, there is a deliberate push for obtaining data of high contextual accuracy but this simultaneously makes maintaining data quality increasingly complex. A similar dichotomy exists with using ICT in countries like India where urban slum areas and smaller towns are still developing the means to conquer digital divides and patchy telecommunication capacities. Tussles over data ownership and agenda-setting roadblocks continue to structure outcomes of data-driven decision making for slum communities on individual and collective scales.

Given these challenges, the central questions of this case study focus on two core issues within the larger spectrum of concerns for data justice. First, issues of accuracy, digitising data and the use of ICT; and second, issues of ownership and translating data-driven knowledge into actionable forms. Through these, this case study engages with the existing discourse on data justice and attempts to bring out real-time challenges of the process of data creation and the potential held in producing actionable data that communities may use in making governance and planning processes more accountable.

The subsequent section elaborates on the central analytical framework used in this study i.e. the 'information value chain' adapted from the work of Richard Heeks and Satyarupa Shekhar (Heeks and Shekhar 2019). Details on the methodology used for developing the study are presented in the following section. Findings detail how barriers and real-time challenges emerging from on-ground practice structure pro-equity development initiatives in order to reflect on the information value chain in practice from the context of developing cities in India. Reflections emerging from the case study analysis are presented in the concluding section with recommendations for a future agenda.

B. Background

Scholars and policymakers have acknowledged that developmental agendas in the contemporary world have moved on from being merely informed to being driven by data (Taylor 2017). Despite this technological overhaul in structuring the governmental agenda, the domain of data and its potency can be circumscribed for many parts of the population. Cross linking, mining and deploying data carries risks of further exclusions and reinforcing inequalities especially for those on the margins. For instance, who gets framed as key perpetrators in the data on air quality in Delhi indicates not just the stakes but also the politics of data representation and speaks of the power inscribed within the collection and presentation of large-scale digital datasets. In the process of datafication, a farmer can be a data point divorced from issues of soil salinity, ground water shortage or crop failure, as s/he seems located within a system but divorced from its structural deprivations. In the process of assigning blame for air quality in the capital, does the responsibility of inner-city industries producing noxious fumes or private motorised vehicular traffic get subsumed in the larger narrative of pollution spikes due to stubble burning by farmers in North India? Who gets penalised in legal proceedings may often flow from the results of these representations; and in this case lead to hefty penalties on the middle- and small-scale farmers who supply Delhi with its grains (Mukerjee and Goswami 2016).

The ability of data to frame narratives or drive agendas speaks of a larger ecosystem and how it interacts with the core data processes of any initiative, project or information system. Those core data processes can be understood as an 'information value chain' (see Figure 1). The information value chain model used by Heeks and Shekhar (2019) visualises the process of datafication with its constituents at upstream, midstream and downstream stages.

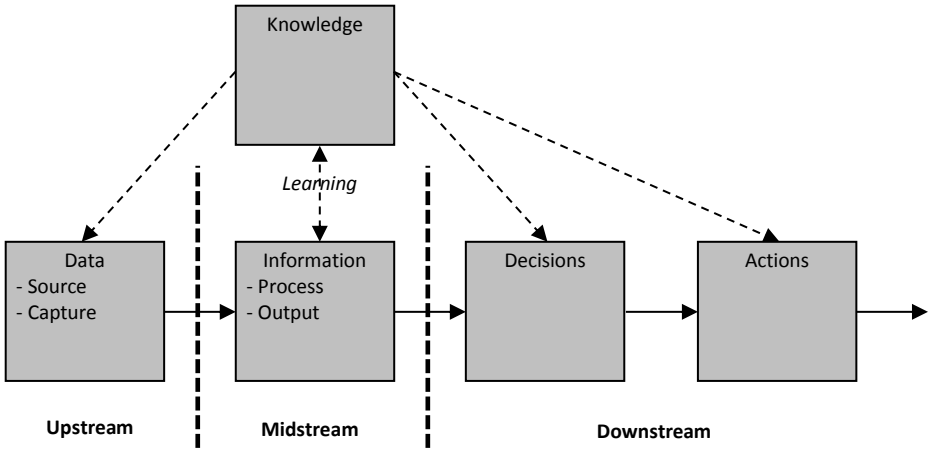


Figure 1. The information value chain (adapted from Heeks & Shekhar 2019)

Larger structural realities and socio-economic vectors sit around these core data processes. They not merely influence information value chains but get consequently shaped by the results and outcomes of information produced through datafication. The entangled nature of data and broader contextual reality thus urges a deeper look into the upstream (e.g. needs, tools and collection processes), midstream (e.g. processing and analysis), and

downstream (e.g. decision making, action plan dissemination and enactment) components of the information value chain. These can therefore be seen to either counter or reinforce discrimination, redress or perpetuate invisibility for those without social privileges, empower communities or sever their rights, etc.

What can data justice mean in such a scenario? Richard Heeks and Jaco Renken define data justice as "the primary ethical standard by which data-related resources, processes and structures are evaluated" (Heeks and Renken 2018:92). They break this down into different dimensions like instrumental data justice (fair use of data and its outcomes), procedural data justice (fair handling and management of data), distributive data justice (fair distribution of data). Threading Amartya Sen's capability approach within this idea imparts agency and emphasises the interlinkages of data/evidence and decision making to improve life opportunities (*ibid.*). An information value chain hence functions and thrives within an ecosystem of social action, interest groups and other external factors which all come to bear on the production, operation and outcome of the value chain.

While the information value chain model (Heeks and Shekhar 2019) provides a central framework for this study, the arguments presented here build upon how this might work in real-time practice and what that could mean when one speaks of data justice. It attempts to foreground the dynamics of heterogeneous experiences of ICT and digital reach, iterative datafication techniques, and limitations in human proficiency for turning data/evidence into action. These are issues that remain true for a wide majority of slum residents in Indian cities and hold an overarching influence in not just the data outcomes but the process of datafication itself. A capability or a choice to shape data outcomes may be highly varied across different locational and socio-economic contexts; not just for individuals but organisations like MHT itself which attempt to strike balances between pro-poor advocacy and producing unbiased data sets.

What this effectively suggests is the urgent need to ensure data use is open, accountable, effective and responsible especially when data intermediaries (organisations working in the zone between the state and the citizens) frame and actualise pro-data equity initiatives. Linnet Taylor notes that, "the framing of data justice … operates at the highest level – where the social contract is shaped and negotiated – and at the most basic, in the practices of everyday … life" (Taylor 2017:11). Given this all-encompassing influence in an increasingly datafying milieu, there remains a significant need and potential for building "ethical paths" (Taylor 2017:2) and embedding a social justice agenda within the larger ecosystem of datafying developmental regimes (Taylor 2017, Heeks and Renken 2018). In the case of action-oriented projects being run for and/or in slum communities, this takes on particular forms and challenges.

Given the digital divide and existing socio-structural hierarchies, data related issues – who collects the data, for what purpose, using what tools for collection, with what level of participation of the community and what availability and openness of data sets, how data is framed or harvested, and what balance exists between privacy and representation and between accuracy and manipulation – all shed light on the multiple intersecting streams of contestation that lie across the information value chain. While discussing Bowker and Star's idea of efficacious systems, Taylor extrapolates that a good, usable system becomes so

convenient that it disappears and only its answers remain (Taylor 2017:6). Similarly, the 'work' in a data-driven world carries a messy underbelly of stakes, interests, costs and controls which catalyse processes even as they remain in the foreground; while the information they hold travels and is harnessed across different contexts or purposes.

While it remains impossible to eliminate the effects of all of these challenges, using a data justice framework allows for bringing self-reflexivity into working with/through data. Hence this study uses the information value chain as a starting point to analyse data-justice-in-practice by reflecting upon interventions and action-oriented projects. The two core issues it identifies attempt to highlight the complexities at two different stages within the information value chain from the perspective of slums and their residents in India (see Figure 2). The first issue is of data accuracy and digitisation of data that pertains to the upstream components of information value chains. The second – focussing on downstream issues – is that of linking data ownership to the community as well as the challenges faced and potential unlocked in harnessing data/evidence.

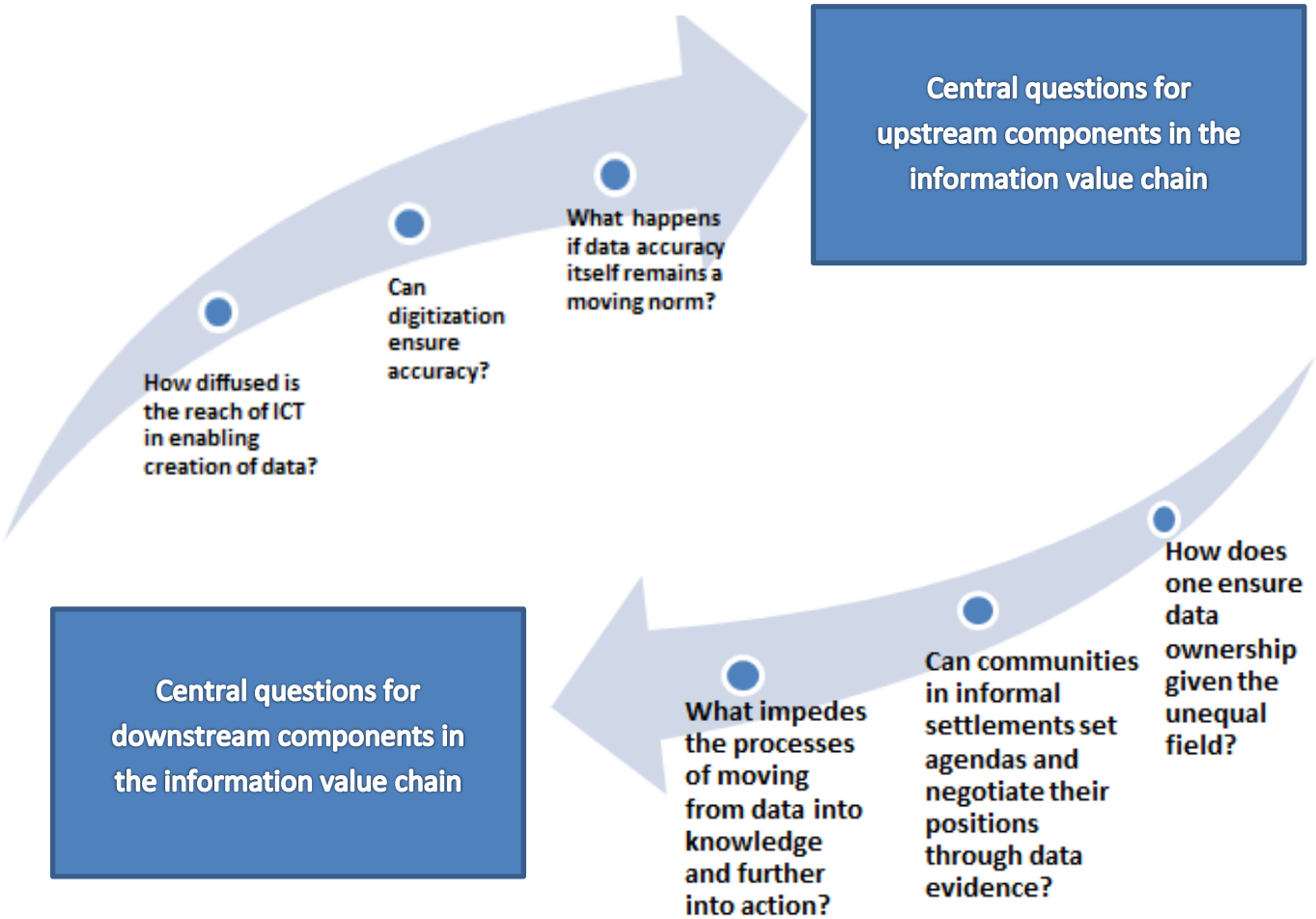


Figure 2. Central focus and key questions of the case study

C. Methods

A combination of secondary and primary research is employed for the purposes of the study. The key framing issues and literature review is carried out through a secondary analysis of some key literature around data justice as a conceptual category and a modular form. In addition, the study is pivoted on a critical evaluation of Mahila Housing Trust's own experience of working with data through examining illustrative examples, experiences and on-ground learning from three key projects:

- **Slum Mapping Exercise in Ranchi (ongoing):** The project was a slum boundaries, infrastructure and service delivery indicator mapping exercise done via community participation in order to produce digitised, updated and geo-tagged maps for 20 informal settlements which have remained largely invisible to date within urban governance in the city. The process was based on translating paper maps made via community collaboration and household-level data collected via mobile apps by trained MHT personnel onto geo-referenced formats for community awareness and mobilisation as well as public advocacy.

- **Energy Audits in Ahmedabad (ongoing):** The Energy Audit project is a community-led energy mapping exercise where women leaders from the community (trained by MHT) in Ahmedabad collect household data about energy expenditure and requirements. Post-analysis, families from informal settlements were provided with customised solutions through microloans, energy-efficient products and suggestions based on their household layout, need and capacity to pay. The collection was carried out by community women through audit tools and paper-based surveys processed and analyzed at MHT with the onus of decision making and action plans on the community itself. The project aimed to ultimately develop the capacities of community-based organisations to monitor and maintain infrastructure involving them in meter reading and the supply and payment of bills.

- **Global Resilience Partnership Challenge (GRP)-supported project building climate resilience for the urban poor (2015-17):** The project titled "Women's Action towards Climate Resilience for Urban Poor" was run in seven cities of South Asia to co-create data and action plans on housing typologies, climate risk and current conditions at the household and slum level. Collected through mobile apps and shared with the community, it allowed for them to create community-level plans and mobilise towards ensuring public accountability.

Primary evidence on challenges and reflections on data justice in practice was collected through qualitative research tools such as unstructured interviews and focus group discussions (FGDs). All primary work was conducted between January-April 2019. The FGDs were targeted towards three different levels of data 'agents' or stakeholders within the organisation, each of whom play vital roles in shaping project design, what kinds of data we seek, what tools are used to handle and secure the data, but also shaping how data is circulated and used.

These levels correspond to: on-ground field team, data management team, and senior management at MHT. The field team is the key data gathering unit and works closely with the community. For the purposes of this study, the field team from MHT's local office in Ranchi in Jharkhand was involved due to the challenging, diverse and wide-ranging project portfolio of ongoing projects at that location. The data management team not only design research tools and survey timelines but perform various upstream and downstream tasks like ensuring training of surveyors, aiming for maintenance of research ethics in the field, handling software and back-end processes, checking and cleaning data, and handling its analysis and presentation. They include a variety of consultants and outside experts as per the need of the project and mostly work out of the head office in Ahmedabad. The team has been moving towards centralising databases through an MIS system for ease of operations and higher consistency.

The third nodal point for FGDs was the senior management who set agendas and steer the work of our projects and organisation at large. These may include but are not limited to those who are in direct contact with donors, policy makers, sector experts, researchers, etc. This level includes those key members from MHT who decide what data may be needed, how the data reports may be presented and shared, and how knowledge that is produced through our work gets disseminated.

D. Findings

D1. 'Fixing them on the maps'

The moving parts of a machine often determine its collective function but to unpack the 'process' of datafication is to grapple with an emergent form whose life and creation remain inherently under flux. What happens if the idea of a norm is no longer static but shifting due to the realities in which it is achieved? How and where does one set a benchmark or fix data points? The question of defining data accuracy presents such dilemmas in the practice of creating data for slums in Indian cities where discontinuity and inconsistency often puncture the idea of achieving accuracy.

In many smaller Indian cities that are still transitioning from predominantly rural to urban forms, like their counterparts elsewhere in the global South, informal settlements due to the insecurity of their tenure and disrupted domesticity either fall off the map or remain hard to capture. They continually mutate or incrementally build in relationship with the externalities they encounter from eviction, heat stress, civil strife to disasters, etc (e.g. Chandran 2019). To capture these marginalised communities and aim for data accuracy through the use of ICT remains equal parts potential and challenge. To use community participation as a means of improving accuracy and create data from the ground up can provide added granularity to GIS maps or drone mapping that may rely solely on remote sensing or official survey maps. However the experience of immersive fieldwork and community involvement also raises concerns over local conflicts, exposure to induced errors, and technological hurdles in remote locations which posit accuracy as not an unattainable but a shifting norm.

For instance, we can take MHT's experience in the ongoing slum mapping exercise in Ranchi undertaken through a complementary process of field-based studies and subsequent digitisation of slum surveys and service mapping via GIS. This illustrates some of the dilemmas in seeking to ensure data justice. Essentially the project was devised as an iterative process of following four stages (see Figure 3). The first stage involved immersive fieldwork and iterative processes to map slum boundaries and enumerate households. While the slum boundary mapping was a collaborative effort between field teams from MHT and community residents, the household-level data was collected by trained field surveyors from the organisation itself. The next stage involved digitising these into geo-tagged formats. To this another layer of data was overlaid by collecting information about household particulars and service levels within the community. In the fourth stage, this information was collated, verified and presented in the form of digitised slum maps for communities which have remained on the margins of state discourses and planning paradigms, in order to initiate public advocacy and to initiate community-led demands upon the state for civic entitlements in informal settlements.

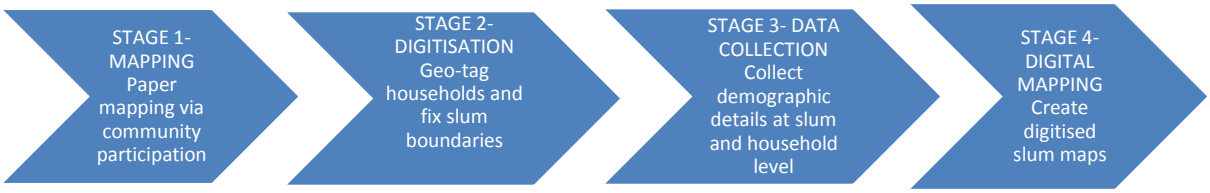


Figure 1. Steps for slum mapping in Ranchi in MHT's ongoing Slum Mapping project

The slum mapping exercise in Ranchi was highly dependent on community knowledge to establish settlement boundaries and service deficiencies in the settlements for enabling service provision and furthering advocacy. While a participatory exercise improved chances of community control over how their reality was captured, this came at a cost where higher chances of inducing manual error, the environment with slums, and attrition rates all shaped outcomes. MHT's Ranchi field team explained that problems arose from a range of issues. Illegal subdividing of houses for rental yields kept changing total numbers of houses for each slum as new doors and temporary partitions often sprang up in subsequent visits to the same slum. Locked houses due to seasonal migration or high incidence of daily wage labour left verification processes reliant on secondary and tertiary sources which often became the neighbours or village-level officers who had little idea about the on-ground situation. Topographical inaccessibility such as open drains and low-level areas which were flood-prone also made coverage of entire settlements and in-person investigation harder especially during monsoons. Organic street networks and fuzzy settlement boundaries due to constantly building and shifting within the settlement made fixing street segmentation and clear boundaries on digitised formats an inherently unfinished business.

In these cases triangulation of data gathering became necessary (for instance, asking respondents, cross-checking with neighbours and via personal observation). But surveyors were often put in decision-making dilemmas, whether their intervention towards achieving accuracy meant inducing bias as a corollary to reducing error. Illegal service connections were observed and often reported by field teams in cases where a true picture at the slum level would benefit from reporting the truth about the household-level picture. Field

officers also noted higher accuracy in household-level data when slum residents perceived that they would be beneficiaries of targeted projects rather than participating in background exercises for advocacy such as the slum mapping exercise. The field teams stated that they operated on a 70-90% confidence level in terms of data accuracy where micro details may only be gathered through using proxy indicators and/or attempting triangulation through asking the respondent, their neighbour and self-observation. 'Fixing them on the maps' and 'making the maps' were hurdles which were constantly being negotiated in the field.

As a data intermediary, working in local contexts also meant dealing with low technological proficiency and testing field conditions. The potential of digitising information about slums through ICT faced high barriers in Ranchi where network issues, limited literacy of field surveyors and respondents as well as geographical inaccessibility produced impediments to creation of data and to translating data into legible and actionable forms. Patchy mobile network connectivity and use of basic devices often led to time lags and poor records of locational accuracy during geo-tagging procedures. For instance household-level surveys were often backlogged due to unavailability of data services in informal settlements in the periphery of the city. It also resulted in delayed updating of survey forms and in extreme cases, mobile device operating system crashes or jumbling of back-end data.

Building staff capacity and maintaining trust with communities was and remains an ongoing challenge. For MHT, building community trust has been a core focus which entails a long drawn-out period of informal field visits, interacting with the community and understanding their needs as well as laying foundations for building community-based organisations in the informal settlements across Ranchi where MHT is working. This initial trust-building process takes around 3-6 months and usually precedes the project or other need-based collecting of data. If absent, it can lead to sources of bias and inaccuracy. From MHT's experience, it is increasingly apparent that data accuracy is directly proportional to levels of mutual trust and intended benefits. Field teams noted that residents of informal settlements shared more accurate data when they have familiarity with and trust the surveyors / organisation.

Women-led teams also encountered hostility and uncomfortable situations at times. Cases were reported of catcalling as well as questioning of women surveyors by male slum residents who were drunk and/or felt threatened by the presence of female surveyors in their settlement collecting data. This forced the organisation to have male members within the team in order to ensure safety. Such involuntary considerations as well as disruptions in the field not only increased the time lag in data collection but also bore financial consequences for the project. A detailed flowchart of the challenges faced at each step is presented in the Appendix.

Experiences from practice urge a rethinking of the idea of data accuracy: not as a static benchmark but a shifting norm structured by the on-ground realities including its emergent milieu and conflicting stakeholder interests especially in areas whose ontology is constantly under flux. While the community felt empowered in terms of participating in creating data and being visible, there still remained discomfort and hesitation in sharing household-level information due to the perceived illegality of their status. Use of ICT facilitated digitisation processes but led to entry level barriers and operational hindrances for local research teams

and the community. Overcoming processual lags and ensuring validity was thus an incremental effort which was enabled through purposive decision making and multiple rounds of cleaning and re-verification. Inducing bias during datafication was wagered against presenting a real-time picture of informal settlements and their needs, while minimising their exposure to punitive action from state and non-state actors.

How context itself shapes the benchmark for data accuracy reveals the contested territory of fixing data while working with information value chains. It reveals that the structural data justice framework holds up a more destabilised and messy world of iterative processes and interactions within action-oriented projects in informal settlements where feedback loops, hierarchies of knowledge and technique, and constant shifts shape more than just the form that data takes.

D2. Data ownership and translating data/evidence into action

Another central focus for data justice discourse has been ensuring that ownership and control of the data created about individuals and communities remains accessible, understandable and actionable for them, in contrast to the intentions for big data. For Taylor (2017), who advocates building ethical paths in a datafying milieu, balancing between integrating the freedom to be seen and represented appropriately with the freedom to be unseen and free from intervention remains a central concern and in practice, a constant challenge. With the marketplace for data expanding drastically – making data one of the most precious and potent commodities in the world today – democratising its access and routes to its use for shaping outcomes has become even more indispensable. Data intermediaries, especially those working for developmental and pro-equity initiatives, need to push for ensuring that communities remain a vital stakeholder in not just generating but owning and using the data generated through them, even if translating data/evidence into actionable forms may prove to be a variegated experience. Community empowerment is pivotal in translating data/evidence into actionable form: merely making data/evidence available to a community does not necessarily translate into desired action.

What are the real possibilities of creating data ownership amongst communities? Additionally can we rethink slum residents as integral parts of information value chains at different stages: from upstream catalysts like research surveyors to data/evidence-driven action agents? Experiences from ongoing projects provide a highly diverse scenario in this regard with every project, location and context defining its own set of hierarchies, hurdles and possibilities.

The energy audit project run by MHT in cities like Ahmedabad provides an interesting example. As part of the project, MHT trains women leaders from slum communities to work as energy auditors in the households within their communities where they audit households on the basis of their energy usage by generating and collating data on electricity bills, consumption and family preferences. The auditing tool they use maps the existing layout of the interior of a dwelling and the light sources and spatial use given many dwellings also double as workspaces for their residents. This mapping helped the women leaders themselves and family members of the households in question to understand how they used their home spaces for different activities. Improvements are then suggested in the

location of light points e.g. splitting the light points to give focused brightness but using lower wattage sources.

Up to this point, MHT has trained a group of nine women across Ahmedabad, who now work as energy auditors for their community. While on the one hand the training empowered community members in creating and analysing their own energy needs, the communities remain dependent on trained experts and architects on second-order issues such as devising layout-based solutions for each household. The process strengthened the community in domains of awareness and decision making at the household level as it prioritised the involvement and needs of the families it aimed to study and benefit. The initial gains on data/evidence-driven action were taking shape through decision making on energy consumption and shifts being made on the level of the household to tackle heat stress, ventilation and low lighting issues as well as improvement in storage capacity and spatial usage within the household unit. By keeping the process and solution localised and indicative rather than directive, the families and community had greater control over the data-driven solutions they wanted to experiment with and invest in. Most of the data collection was still done through paper and usage of electrical equipment such as voltmeters required significant external support. Local identities, politics and socio-economic conditions still held a significant impact on what action could stem from the awareness that is generated.

Additionally these energy auditors also introduce low-cost, energy-efficient solutions and products manufactured by MHT through subsidised means. Through its network of women energy auditors, MHT has sold sustainable energy products such as LED lights, smokeless cook stoves, solar lanterns and solar lighting, and cooling systems to more than 27,500 slum families across 10 cities in India. Yet this network made the women in the community answerable and accountable to the organisation. The data being collected was held in a central repository at MHT which had the final say on its usage. The process of democratising all the way along the information value chain hence remains far from over and requires more work. Even while one domain may acquire more openness or robustness, others produced new hierarchies in terms of how data was processed or where it was sought to be deployed.

While the energy audit process trains community women to recognise and devise solutions for rationalising their energy needs at household and neighbourhood level, the umbrella project for women-led climate resilience building under the Global Resilience Partnership (2015-17) pushed the agenda of data/evidence-driven civic participation further along the road to achieving data justice. Based on evaluating, harnessing and improving community capacity in seven cities across South Asia[2], the project, with its process methodology detailed in Figure 4, was designed to shift the sustainability quotient and driving gear of the datafication and action planning process into the hands of the community. The project was conceptualised to ensure:

[2] Ahmedabad, Bhopal, Bhubaneshwar, Jaipur and Ranchi in India, Kathmandu in Nepal, and Dhaka in Bangladesh.

If the urban poor are provided with the requisite knowledge to undertake vulnerability and risk assessments and are equipped with accessible resilient technologies, they will be able to devise and implement locally relevant and pro-poor climate resilient solutions. If the poor are empowered to implement their own resilience plans, and the institutional mechanisms representing their voices are in place, they will be able to better influence city planning and governance on pro-poor adaptation and resilience action (MHT 2017:ii)

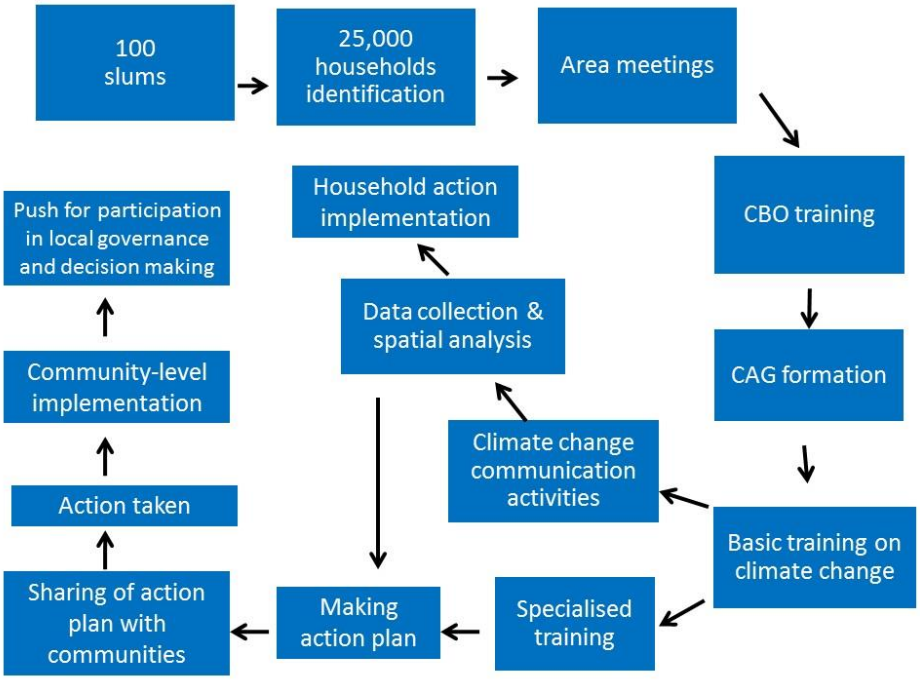


Figure 4. Methodology for carrying out MHT's climate resilience building programme at community level under GRP

Premised on building community-based organisations (CBOs) and their active proponents called community action groups (CAGs) at the slum level, community participation was central in each stage from collecting data to training to negotiation of action agendas by interacting with local municipal authorities. The GRP project allowed community action groups to form community-based action plans for each site according to their own priorities against the identified risks of water inundation, flooding, heat stress and vector-borne diseases. Use of data/evidence in setting priorities – not just for local decision making but for improving governance – pivoted on motivating and externally supporting community mobilisation based on data/evidence at the slum level which was being generated for the first time and pushing them to urge for inclusivity in local governance.

Some of the success towards this goal included community leaders from slums in Ahmedabad being made stakeholders in devising municipal authority Heat Action Plans; in Bhopal they were invited by the city municipality's Malaria Department to undertake vector surveillance in other slums; whereas in Ranchi community leaders managed to push for including rainwater harvesting incentives in local implementation of the national-level housing policy, Pradhan Mantri Awas Yojana. The women in the community were slowly trained to identify and understand technical terms and conceptual categories like climate

change, risk, heat stress, vector-borne diseases, green roof – ideas that will permeate into the memory and vocabulary of the community over a longer, more sustained period of time. New initiatives such as climate *saathis* (partners) and child doctors were instituted to involve adolescents within the community.

Despite the positive response from the community and growing platforms for them to interact in local governance, many challenges still remain. The GRP programme provided the first push towards participatory processes where slum communities could interact with government to demand entitlements or better provisions but requires scaling for this effect to bring a significant change in how communication channels in local contexts are structured. The research design for a major portion of the household data collection exercise was undertaken by MHT's research teams who collected, collated and analysed the datasets produced cross multiple sites. A significant amount of training and external support remained necessary throughout and after the project in measuring and collating temperature variations, or maintenance of equipment for installed green roofs and modular roofs.

While the community members played a participatory role in producing data and deciding what forms it could take while translating evidence into community-level action plans as was envisioned in the initial project design, their role in taking it to the level of participation in local governance remained dependent on MHT's support, from organising women leaders to fixing meetings or creating platforms for engagement. On different occasions, interactions between stakeholders at meetings with local authorities had to be facilitated by MHT. Often the agenda setting and modes of intervening in local decision making were devised at the organisational level and diffused into community meetings, later being picked up by them and pursued depending on their feasibility.

Additionally while data collection and action plans still possessed an active component of co-creation at the community level, the downstream channel of data dissemination was narrower than desired due to logistical issues and privacy concerns. As a result, collated datasets were managed and remained with MHT itself even if their broad results were used by the community members. While data justice aims to close the loop on accessibility and dissemination of data/evidence for the communities who may be represented through it, deficits in terms of organisational planning about how this could be made available to them as well as prevalence of digital illiteracy amongst slum residents hampered the cause of open access and free use of post-analysis trends and evidence on risk and resilience capacity.

As illustrative examples, both the energy audit and GRP project present different scales and modes of how slum residents generated data/evidence and participated in decision making. For the energy audits, data creation and its concerted use was squarely at household and community level. This made the transition to action comparatively easier. The local dynamics and household considerations determined interactions and responses, where dependence on organisational support for technical proficiency and solutions remained considerable. In this case, the field surveyor, data analyst and decision maker are all from and within the community – a relatively rare scenario. This was enabled through empowering and training the community. While these required high MHT involvement in

the primary stages, it eased community reach and closing of the loop from knowledge to action.

With GRP, the community's dependence on MHT for understanding issues, building and mobilising the community, and catalysing their interactions with city-level decision makers remained significant. Results for data analysis were shared with the community but the collated and analysed databases were held at the organisation level. Even while these remained accessible for community members, making them practically usable has remained challenging due their digitised form. Demonstrative results from the analysed data such as results from water testing or vector-eradication drives was what the community largely relied on to understand and prioritise their resilience strategies. As complexity of data collection, analysis and decision-making process increases, community participation and ownership becomes harder to ensure yet necessary for enabling free and fair use of data; i.e. necessary for data justice.

E. Discussion and Conclusions

The projects and experiences from the field provide insights on both upstream and downstream issues of information value chains. The slum mapping exercise in Ranchi brings in contextual learnings relevant to similar locations undergoing rural-urban transitions and seeing high rates of attrition and incremental building in informal settlements. Their location and topographical conditions produce challenges of access for fieldworkers and for fixing them in digitised and geo-tagged formats with cohesiveness and complete accuracy. In addition, high political stakes and tenure insecurity drive communities to make conflicting or inconsistent claims that push field researchers to triangulate data and make choices between correcting or inducing bias. The upstream stages of information value chains remain extremely entangled in practice i.e. moving back and forth due to procedural and ethical dilemmas of fixing accuracy. Further, digital divides and limited diffusion of ICT have meant that it often creates hurdles for both researchers and the community in implementing research as well as translating information into action. Hurdles such as low network connectivity, limited ICT diffusion and low digital literacy change the discursive domain of what can be achieved in the course of enumerating and mapping slums.

On the other hand, experiences from practice in the energy audit programme and GRP project help reflect on how issues of data ownership and challenges to data-driven civic engagement shape the downstream effects of information value chain operations. While the energy audit was at the settlement level and improved chances of keeping the information value chain operation grounded and community-led, on macro, multi-city scales like those in GRP this became harder. Slum communities in Ahmedabad needed support and handholding in enabling action through data/evidence-driven negotiations. Socio-economic conditions, gender biases, literacy levels, and digital illiteracy continue to be entry barriers in pushing for desired outcomes. Creating chances for data-driven civic engagement required ongoing organisational support and facilitation. Data ownership in the case of GRP at least remained circumscribed to the domains of intermediaries, research organisations or state actors. While this process of using data/evidence-driven action is getting

democratised, representation of macro data sets, rules for access and maintenance of e-repositories remains harder to decentralise.

Improving MHT's own data and project documentation has also been a challenge that the organisation continues to strive for. Building consistency in our data collection and handling will allow for better preparedness towards improving channels for access. In addition, another key learning from ensuring data justice in practice is to build the familiarity and capacity of working with digital formats for those within the organisation as well as the communities. Developing technical knowhow and capacity building by intermediaries like MHT which work at grassroots level may be a step towards improve the ability of slum communities to effectively understand, handle and yield their data in future. Another learning has been towards developing data collection tools that incorporate multiple formats suitable to support triangulation. For instance, survey questionnaires for respondents that incorporate input media such as audio and taking geo-tagged photographs of locations improve data cleaning processes and accuracy levels.

While the promise of data justice could restructure hierarchies, the new forms of dependencies being created also need attention. For informal settlements in the global South not just digital literacy but literacy more generally remain an unfinished goal; yet a basic foundation for data justice in a digitising world. ICTs can play a role to democratise and decentralise datafication processes but only if we move towards fairness of use and representation, and enable a rights-based discourse that balances the need to be seen with the right to determine how and the space to refuse. A contextually-grounded evaluation of the work of information value chains allows us to consider the domain of possibilities against contours of impediments, both of which remain inherent in the processes of a data-driven environment. The discussion on data justice in MHT's case study runs across multiple registers. First, the case study allows us to engage with logistical or ethical concerns that emerge in practice such as dealing with concerns regarding efficacy versus limits of ensuring accuracy, and working in technologically-deficient locations or with conflicting interests. Second, the analysis of the projects illustrating data/evidence-driven engagement highlights how data can intervene in the lives led by the urban poor and in the challenges that continue to plague them.

A future research agenda towards building data justice into practice needs to focus on better understanding of capacity-building interventions with communities; particularly identifying strategies for them to actively engage with and own the process of creating actionable data for change. This requires them to be not merely providers of data but stakeholders in the process of envisioning data-driven civic engagement. Another key focus for future research will be scaling decentralised data initiatives. In small-scale or community-based projects such as energy auditing, there was considerable evidence to suggest that the envelope on data ownership and subsequent decision making was being pushed. However, this remained a challenge for macro-projects like GRP. Building grounded frameworks and models for scaling interventions that aim for data justice in practice remains a necessary next step for a just and fair future.

References

Arora, P. (2019*) Benign Dataveillance - the New Kind of Democracy? Examining the Emerging Datafied Governance Systems in India and China*, Communicative Figurations, Working Paper 24. ZeMKI, Centre for Media, Communication and Information Research, University of Bremen.

Banerjee, S. (2017) It's time to disentangle the complex Aadhaar Debate, *The Wire*, 20 Apr. https://thewire.in/government/aadhar-privacy-analysis

Chandran, R. (2019) More than 11 million face eviction, displacement as India booms, *Reuters*, 9 Jun. https://in.reuters.com/article/india-landrights-eviction/more-than-11-million-face-eviction-displacement-as-india-booms-idINKCN1RL2VL

Ghertner, A. (2015) *Rule by Aesthetics: World Class City Making in Millennial Delhi*. OUP, New York.

Heeks, R. & Renken, J. (2018) Data justice for development: what would it mean? *Information Development*, 34(1), 90-102.

Heeks, R. & Shekhar, S. (2019) Datafication, development and marginalised urban communities: an applied data justice framework, *Information, Communication & Society*, 22(7), 992-1011.

MHT (2017) *Women's Action towards Climate Resilience for Urban Poor in South Asia*, Mahila Housing Trust, Ahmedabad

Mukerjee, P. & Goswami, S. (2016) Delhi air pollution: identifying perpetrators and fixing responsibility, *DownToEarth*, 7 Nov. https://www.downtoearth.org.in/news/air/delhi-air-pollution-identifying-perpetrators-and-fixing-responsibility-56240

Roy, A. (2009) Why India cannot plan its cities: informality, insurgence and the idiom of urbanization. *Planning Theory*, 8(1), 76-87.

Scott, J. (1998) *Seeing Like a State: How Certain Schemes to Improve the Human Condition have Failed*. Yale University Press, New Haven, CT.

Taylor, L. (2017) What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, 1-14.

UN. (2016) *Urbanization and Development: Emerging Futures, World Cities Report 2016.* UN-HABITAT, Nairobi.

UN. (2018) *2018 Revision of World Urbanization Prospects.* UNDESA, New York.

Warner, G. (2013) In Kenya, using tech to put an 'invisible' slum on the map, *Parallels*, 17 Jul. https://www.npr.org/sections/parallels/2013/07/17/202656235/in-kenya-using-tech-to-put-an-invisible-slum-on-the-map

World Bank. (2011) *Guide to Climate Change Adaptation in Cities*. World Bank, Washington, DC.

Acknowledgements

The "Urban Data, Inequality and Justice in the Global South" case studies form part of a Senior Research Fellowship funded by the University of Manchester's Sustainable Consumption Institute with additional financial support from Canada's International Development Research Centre. We wish to thank all the members from field teams and local offices who work towards improving our reach towards the community and ways of knowing them through data and relationships we create in improving the habitat of non-networked slums. But most of all, we remain grateful towards the residents, communities and women from informal settlements across India who continue to trust us and remain

enthused about actively shaping the world they inhabit – through data, action and their unwavering resilience.

About the Authors

Bijal Brahmbhatt is currently the Director of Mahila Housing Sewa Trust (MHT). A civil engineer by training and having expertise in habitat improvement, community development, housing finance and urban climate change resilience, she oversees MHT's operations at the national level. She has proven experience in conceptualising, planning, managing and providing support for women-led climate resilience programmes across India. Her professional experience has focused on a range of poverty alleviation issues, particularly with women, entrepreneurship, slum upgradation, water and sanitation, housing and urban planning, housing finance, housing technology and renewable energy.

Siraz Hirani is presently working as Senior Programme Management Specialist at Gujarat Mahila Housing Sewa Trust. He is a senior professional with over 16 years of strategic and operational experience with international development organisations including 8 years in leadership positions at global and national level. He holds an MSc in Disaster Mitigation, MSc in Life Science and MBA with specialisation in Finance. He is an alumnus of the Harvard Kennedy School with work experience in multi-cultural and challenging countries like India, Bangladesh, Afghanistan, Tajikistan and Myanmar. He has proven capability in strategic planning, programme management, policy development, change management, fund raising, forging partnerships and talent management. His sectoral experience includes disaster risk management, humanitarian aid, habitat development, water and sanitation, health, and livelihoods and governance.

Neha Lal is Senior Research Associate at Gujarat Mahila Housing Sewa Housing Trust and works on research, documentation and capacity building at MHT. She holds a Masters and MPhil in Sociology from Delhi School of Economics, University of Delhi and works on the intersections of issues of housing, citizenship, urban governance and planning. Her MPhil thesis looked at the changing terms of citizenship and rights of belonging for urban poor in millennial Delhi. She was also an Urban Fellow at the Indian Institute of Human Settlements, Bangalore. Neha has lived in multiple Indian cities, working on issues of housing rights, governance, water and sanitation and urban infrastructure, contemporary history and urban ecology.

Bhumika Chauhan is a postgraduate in ICT, Agriculture and Rural Development from Dhirubhai Ambani Institute of Information Communication Technology, Gandhinagar, India. At MHT, she has been working on developing research and documentation tools, GIS portals, a vulnerability calculator and community kiosk, and automated indoor hydroponics systems that focus on providing technological solutions to poor communities. With multiple years of experience in IT project management, GIS and programming languages, Bhumika's work focuses on designing and promoting innovative technological solutions for poor communities living in informal settlements in order to mitigate their exposure risk, enable knowledge transfer and push for political advocacy.

Appendix: Slum Mapping Steps and Challenges

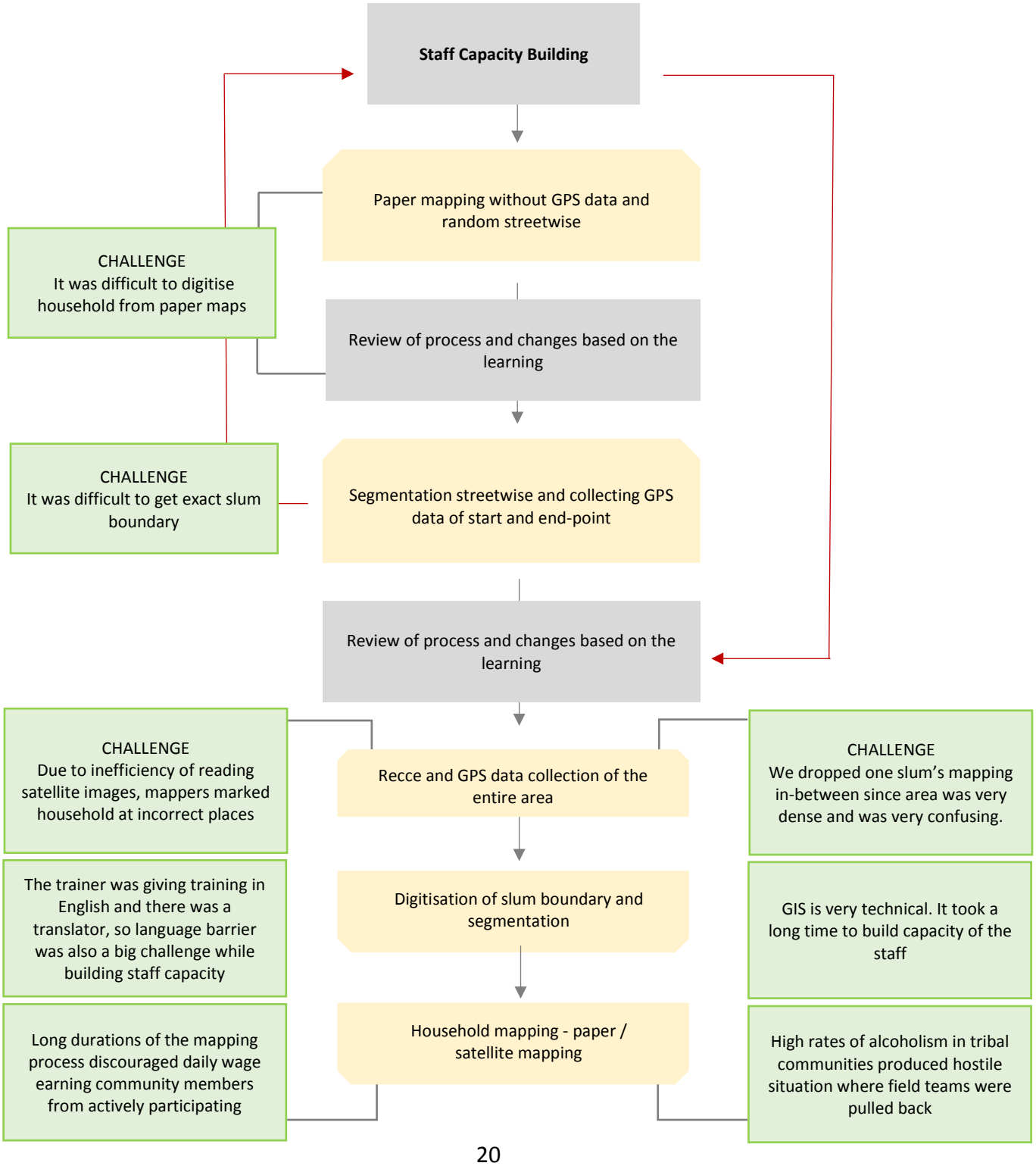
Stage 01
MAPPING
Paper mapping via community participation

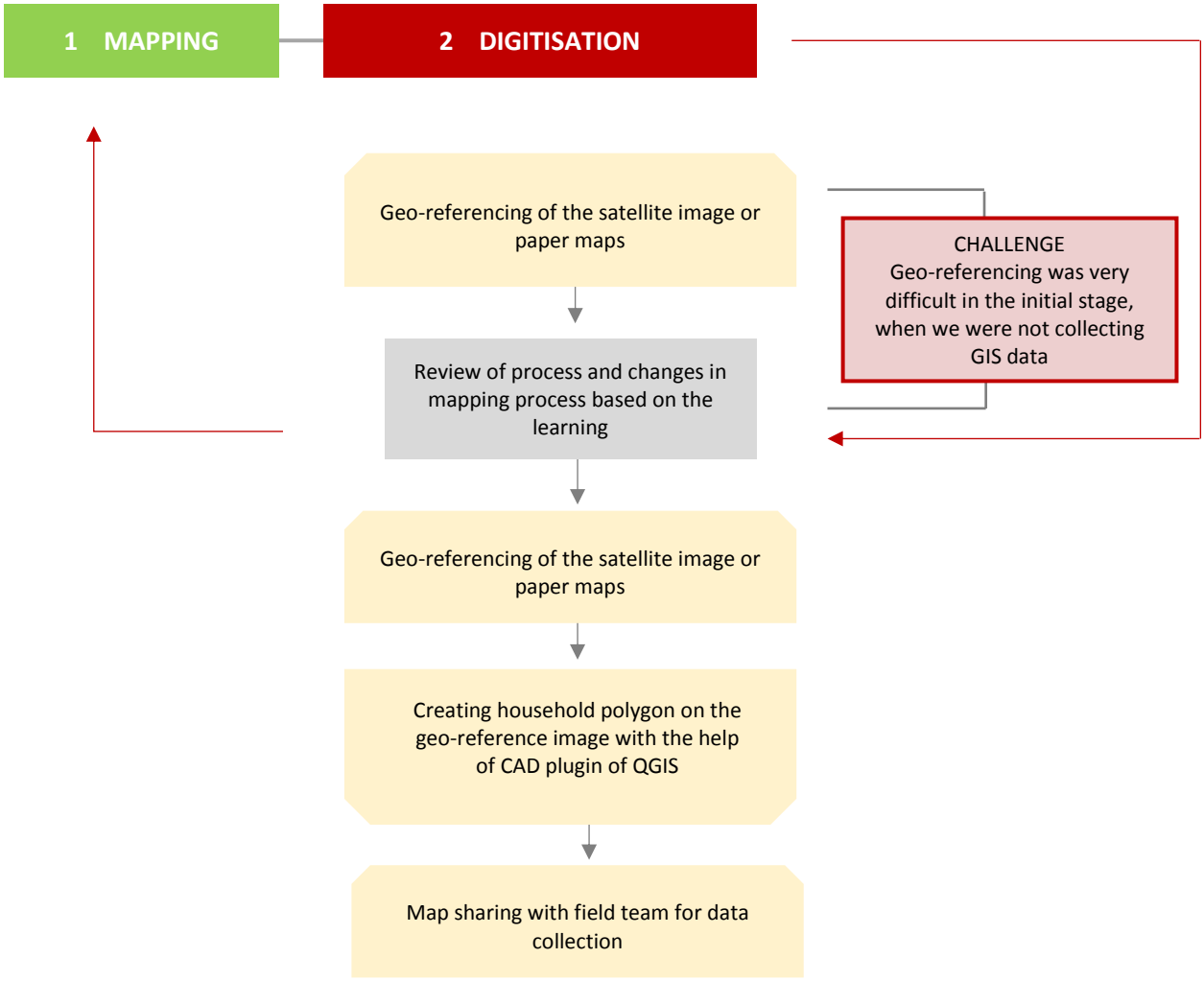
Stage 02
DIGITISATION
Geo-tag households and fix slum boundaries

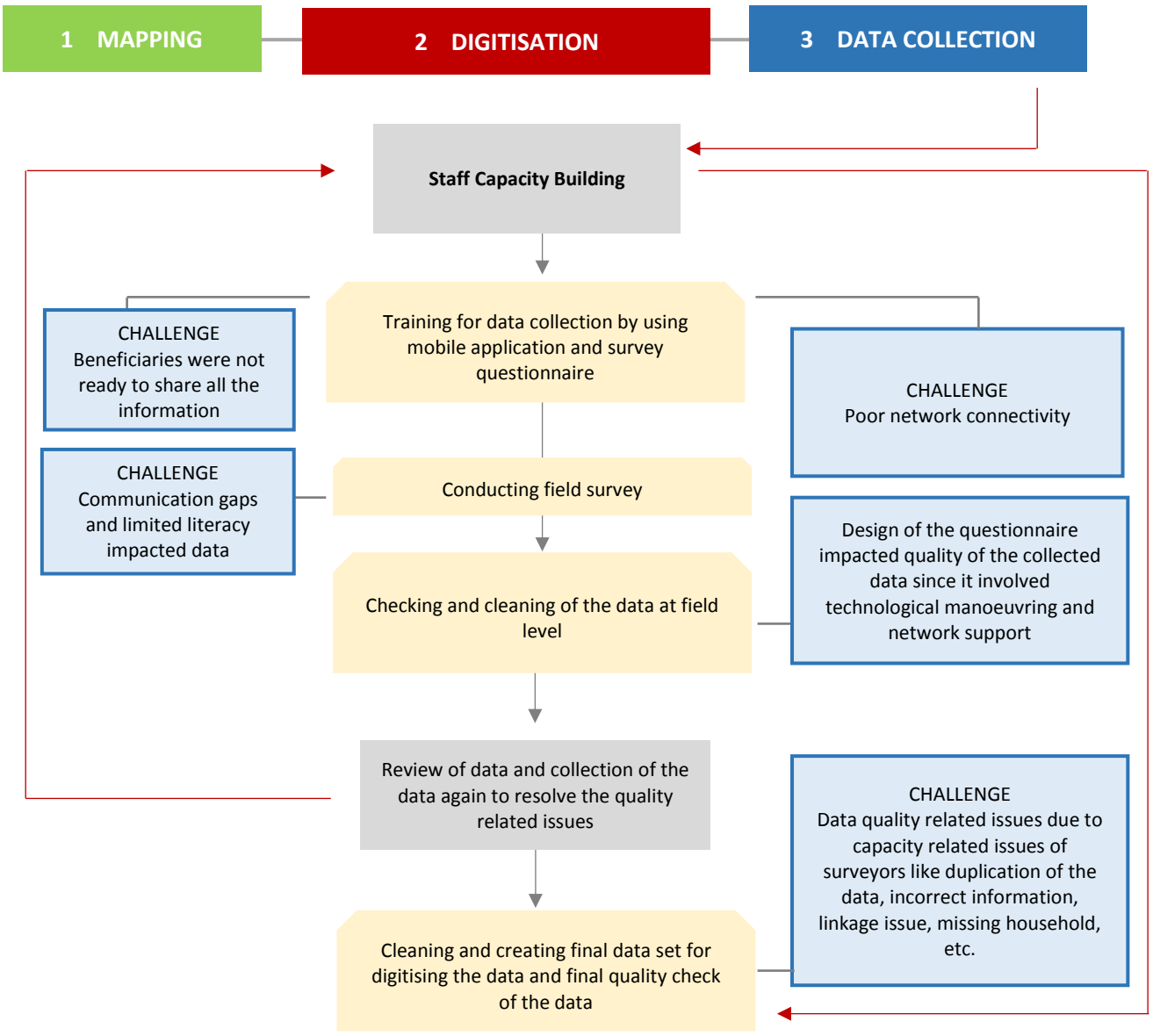
Stage 03
DATA COLLECTION
Collect demographic details at slum and household level

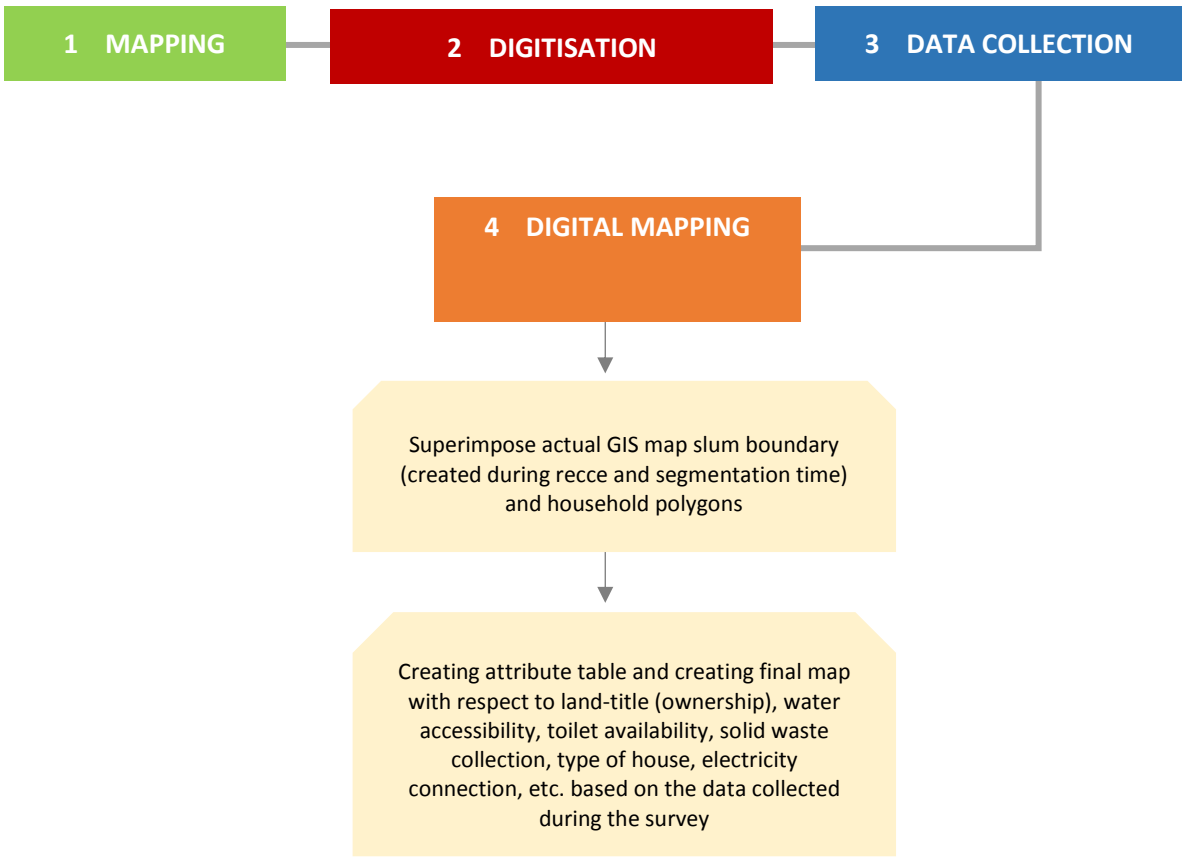
Stage 04
DIGITAL MAPPING
Create digitised slum maps

1 MAPPING




1 MAPPING
2 DIGITISATION
Geo-referencing of the satellite image or paper maps
CHALLENGE
Geo-referencing was very difficult in the initial stage, when we were not collecting GIS data
Review of process and changes in mapping process based on the learning
Geo-referencing of the satellite image or paper maps
Creating household polygon on the geo-reference image with the help of CAD plugin of QGIS
Map sharing with field team for data collection


1 MAPPING
2 DIGITISATION
3 DATA COLLECTION
Staff Capacity Building
Training for data collection by using mobile application and survey questionnaire
CHALLENGE
Beneficiaries were not ready to share all the information
CHALLENGE
Poor network connectivity
Conducting field survey
CHALLENGE
Communication gaps and limited literacy impacted data
Checking and cleaning of the data at field level
Design of the questionnaire impacted quality of the collected data since it involved technological manoeuvring and network support
Review of data and collection of the data again to resolve the quality related issues
CHALLENGE
Data quality related issues due to capacity related issues of surveyors like duplication of the data, incorrect information, linkage issue, missing household, etc.
Cleaning and creating final data set for digitising the data and final quality check of the data


1 MAPPING
2 DIGITISATION
3 DATA COLLECTION
4 DIGITAL MAPPING
Superimpose actual GIS map slum boundary (created during recce and segmentation time) and household polygons
Creating attribute table and creating final map with respect to land-title (ownership), water accessibility, toilet availability, solid waste collection, type of house, electricity connection, etc. based on the data collected during the survey