The Challenges of Using Data in Urban Planning in Pakistan

Nabeel Shakeel Ahmed*
York university, Canada

Abstract

As technology adoption and urbanization accelerate in Pakistan, there is an opportunity for urban planners to adopt new tools and enhance their skillset. While data collection and the use of geographic information system (GIS) tools are common in Pakistan already, little is known about how these are used in the everyday practice of urban planners, or the extent to which planning practice has incorporated innovative methods. Based on a series of interviews with planning practitioners and policymakers, this article presents research on the use of data in urban planning in Pakistan. It begins with an overview of the use, opportunities, and limitations of data in urban planning. The article then shares interview findings and identifies a number of challenges to the use of data in planning, classified into challenges regarding data access and reliability, data and digital literacy, and institutional challenges. As the research scope is limited, the article makes only tentative conclusions about the possible underlying reasons of these challenges that require further investigation and validation.

Keywords: data; evidence-based planning; urban planning; smart cities; geographic information systems

Introduction

Urban planning has always been linked to the application of scientific knowledge, seen as 'rational social action' (Escobar, p.148, as cited in Sachs, 2010). Faludi and Waterhout (2006) cite the survey carried out under the command of William the Conqueror in 1086 as an early example; the most influential modern exponent of this tool was Patrick Geddes, who declared "survey before plan" in the early 20th century (Davoudi, 2006).

In the 21st century, urban planning has become increasingly reliant on the collection and analysis of data, and new tools such as sensors, mobile applications, geographic information systems, and machine learning systems have become commonplace around the world. These new

* Corresponding Author: nabeel.shakeel@gmail.com

techniques and technologies are part of a broader wave of digital transformation underway in city governments and are popularly marketed as smart cities. Much of the debate around smart cities and the use of data and technology has revolved around the global North, despite the rapid evolution of urban governance and practice in the global South. This transition is dynamic but has gone almost unnoticed; there are several gaps in our understanding of how emerging technologies are introduced, harnessed, implemented and impact urban areas in the global South. This article addresses one of those gaps and investigates the use of data in urban planning in Pakistan.

A key theme of this paper is that the adoption of data and technology presents opportunities to improve urban development in Pakistan. Directly, data is critical for the effective operation of municipal services and can help address longstanding inefficiencies. It is also necessary for good planning: Yuen and Choi (2012) have highlighted that a lack of data and information act as barriers to effective spatial planning and consistent policy development and implementation. At another level, technology can help strengthen local democracy by improving transparency and enabling citizen-led innovations that use open data.

This is not meant to reproduce a modernist discourse. This article responds to Anwar's call to "imagine planning practice through alternative lenses that are not truncated by consigning futures to the limited imaginings of developmentalist or technocratic interventions" (2012, p. 96). Data and technology present risks that planners and policymakers in Pakistan need to pay attention to. Technology can be used to further exert top-down control, for instance through surveillance, and data can perpetuate discrimination.

This introduction frames the article by describing the key research questions and methodology, and provides an overview of the role that data and technology can play in urban planning. The following section shares research findings regarding the use of data in planning in Pakistan. This is followed by a section on the challenges to the use of data and the limitations of urban planning in Pakistan.

Research Questions and Methodology

The primary research question of this article is: *What are the challenges facing the use of data by urban planners and policymakers in Pakistan?*

The article seeks to build on an emerging area of inquiry around the use of data, digital tools and information technologies in planning in Pakistan, as well as the global South more generally. A crosscutting theme of this article is to introduce the politics and governance of data and technology in cities.

Nine semi-structured interviews were conducted for this study, with additional comments from providing significant details and the perspectives of key decision-makers regarding the questions identified above. Seven participants were current urban planning practitioners; two others were senior provincial public servants; and another was a private data services provider. Participants were selected using a purposive sampling approach, on the basis of expert knowledge, varying perspectives, and availability. These interviews were supplemented with dozens of informal conversations with public servants and development sector professionals.

The limitation of such an approach is that it is susceptible to biases and limits the broader applicability of the research findings. The positionality of the author also affects both chosen methods and analysis; as a member of the diaspora, my understanding is limited. The article also does not pay attention to the questions of class, gender and climate change, even as the use of digital tools in cities is salient for all three and they merit further study.

Data, technology, and urban planning

After an extended period of relying on the Geddesian approach of survey, analysis, and design and the so-called golden age of land use planning, a decisive shift occurred when urban development accelerated in the 1950s (Hall, 2014). This resulted in 'the systems revolution', where planning went from an art to a science; from conceptualizing planning as the creation of blueprints to following a process and optimizing methods (ibid). At the same time, the field of geography was going through what is now referred to as a 'quantitative revolution'. This was an attempt to transform the field into a scientific discipline like physics, using positivist approaches to validate knowledge production (Kwan & Schwanen, 2009). Mathematical and statistical approaches gained credibility and consensus among a range of geographers, including human and economic geographers (Burton, 1963). This revolution was part of a surge of interest in scientific planning as the 20th century progressed, as computing grew ever more

powerful and positivist approaches dominated in policy and management fields (Batty, 1994). Transportation planning played a critical role in propagating this new paradigm, with a number of successful computer applications (Hall, 2014).

This period reached a turning point, with the development of large-scale computerized land use models in the 1960s, such as a model of Pittsburgh that predicted the location of urban activities based on spatial interactions and demographics. This modeling movement seemed to fizzle out quickly, however, as many models were never completed or failed to meet expectations, criticized as "poorly documented monster programs executed on mainframe computers and understood only by their authors" (Wegener, 1994, p.25). Modeling and other technologies continued to receive attention, but the scale of this work shifted from transforming the city to improving its management and operations (Batty, 1994). Urban information systems and geographic information systems in particular emerged as significant areas of development.

With dramatic increases in computing power and the rise of the Internet in particular, information communication technologies (ICTs) have transformed urban planning and management (along with the rest of the world) over the last few decades. Real-time communications open up new possibilities for rapid decision-making, and the technological limitations to capturing, analyzing, storing and recalling data have long been superseded. Data-driven approaches to management and administration are now standard in municipal governments around the world (Townsend, 2015).

The growth in data collection and storage has led to the emergence of big data, which has been defined as "the information asset characterised by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value." (De Mauro et al, 2016).

A wide range of applications have been identified and tested for the new science of cities (Batty, 2013). In the framework provided by Zheng et al. (2014), there are applications at four levels:

1. Urban sensing and data acquisition,
2. Urban data management,
3. Urban data analytics, and

4. Service provision.

Thakuriah et al. (2017) identify the following:

1. Classical applications such as understanding transportation flows and housing dynamics,

2. Complex systems analysis such as the large-scale models discussed earlier,

3. Empirical research such as that on sustainability and socio-economic disparities, and

4. Remote and collaborative sensing.

There is a distinction between the technologies that are being used (for example, remote sensing and GIS tools such as ArcGIS and data processing tools such as the Hadoop framework) and the applications of those technologies (such as analyzing past performance, forecasting future needs, optimizing operations, and facilitating management).

The evolution of a science of cities demonstrates that the application of data is hardly new in urban planning. What is new is the sheer quantity and type of data now available, as well as the increasing sophistication of tools to analyze it.

Use of data in urban planning in Pakistan

This section shares use of data identified by interviewees: to carry out analysis, identify priorities, prepare development plans, understand densification and urbanization trends, and enhance intelligent transportation systems.

"You can't plan anything if you can't measure it." (Interview, May 30, 2018)

Every interviewee testified to the centrality of data in planning. Urban planners use it for scenario building, supporting decision systems, and developing policy positions. They collect both demographic data as well as spatial data, where they can get access to it. Where the data is not available, they have to conduct costly and time-consuming primary surveys.

An example is the planning that was conducted for Fort Munro, a hill station being developed as a tourist resort (The Express Tribune, 2015). Due

to a lack of data, planners have had to collect primary data on the indicators identified by the government.

At the regional scale, urban planning institutions such as the Urban Unit, a private sector company fully owned by the Government of Punjab, use data to develop planning documents such as the Punjab Spatial Strategy, using both primary and secondary sources. The secondary sources include the Pakistan Social and Living Measurement Survey (PSLM), Multiple Indicator Cluster Survey (MICS), Labor Force Survey (LFS), Punjab Development Statistics (PDS), and other reports prepared by think tanks such as the Punjab Economic Research Institute. For most of these reports, micro-data was not available and had to be purchased from the Bureau of Statistics. Primary research was conducted in the form of a Census of Manufacturing Industries, a previous census of brick kilns, remote sensing and modeling to identify road lengths and transportation infrastructure, remote sensing and field visits to identify administrative boundaries, and digitized Survey of Pakistan maps.

This data has been used for identifying growth corridors, industrial estates, transport networks, and the most desirable connections between these locations. This helps identify areas that need to be protected or preserved, areas that need greater investment, and the nature and distribution of spatial disparities and environmental challenges.

To understand the use of data by planners, we have to explore the broader policy environment in which they practice. A common theme emerging in interviews was that there is plenty of data, but in practice, its use is limited by ignorance and the lack of importance that policymakers place in it. In addition, even if there were sufficient awareness of the value of data, the impact of this use would be limited by a lack of capacity to conduct data analysis, as well as the distortion of data for political reasons.

Several responses pointed to a general lack of awareness of data, as well as the choice of which data to use. For example, an urban planner shared that a common concern expressed about urbanization in Punjab is that land is being converted from agricultural to non-agricultural uses, which could lead to food insecurity. In reality, analysis by the Urban Unit shows that less than two percent of the land has been converted, so this concern may be overblown. In addition, there can sometimes be an excessive focus on gathering micro-level data instead of looking at the macro perspective.

The issue of the use of data was highlighted when I interviewed a senior bureaucrat as well as a junior public servant at the Planning and Development Department. The senior official seemed almost offended by the question of whether data was used. He claimed that it was used frequently, and that even the assumptions they made were based on data. He shared that there was sufficient and reliable historical data for a number of areas. For example, the government can estimate the demand for education at the district level. Land and vehicle records are kept and are reasonably reliable. That said, the senior official also stated that whether the data is used well or not depends on how different projects are being managed and if they are allocated appropriate resources. It is worth noting that while I was waiting to interview this senior official, I spent several hours in an office for planning officers of a specific department, which was only partially occupied and where junior officers were chatting for nearly an hour as they waited for their section chief to return from an interdepartmental meeting

Another phenomenon, according to one of my interviewees, is that government officials hide data because they think they will get in trouble if the truth is revealed. He shared a recent incident where the Supreme Court of Pakistan was investigated water supply and infrastructure in Karachi. Upon finding that there was very little actual data available, the court admonished the officials responsible for not sharing data. This demonstrates that even when data may be available, it may be hidden and left unused as a protective strategy.

It is already becoming clearer that ensuring that data is used is one challenge; ensuring that data is used *effectively* is entirely another challenge.

It is clear that, in contrast to the scientific view of data as a neutral construct, public servants and politicians in Pakistan both recognize that data is inherently political. Politicians use data when they want to ask for funds for pet development projects, or when they want to criticize their competition. An interviewee shared that information is used as a tool, selectively and in order to justify different points of view and for political purposes. Data is not shared openly because of what it might reveal about the performance of the government, and interpretations are hotly contested. Since data is not standardized or defined consistently, there can be multiple interpretations of the same information, leading to confusion.

As described earlier, interviewees considered the use of data intrinsic to their work. When describing the opportunities of using data in planning, several interviewees mentioned that they could improve the quality of their work (such as forecasts) if they had better data. One planner spoke about the difficulty of estimating the pace, scale and geography of densification in cities. He used satellite data and average household sizes and land uses to come up with an evacuation plan, but felt that it was not as detailed as it needed to be. A number of them spoke about the value of improving the speed and efficiency of planning and execution.

When asked about the opportunities of big data in particular, interviewees identified the ability to extract fine-grained insights and accelerate decision science. One interviewee spoke about tracking people using mobile phones to identify travel patterns and peak shopping times as an example of the new level of granularity that could be achieved. One interviewee mentioned the ability to better gather environmental data. Another shared an idea to use footage from surveillance cameras to gauge flooding in different parts of the city during the monsoon.

Transportation planning is often seen as low-hanging fruit to apply big data to. One interviewee spoke at length about how big data could be used to enhance the intelligent transportation systems being used by the Punjab Masstransit Authority for the bus rapid transit systems and upcoming Orange Line metro rail service. The existing Lahore Metrobus is already managed through a command and control center that relies on a network of 430 cameras, loop sensors, and other sensing instruments, enabling bus priority, precision docking, and automated fare collection (Interview, June 1).

Another interviewee spoke about an initiative to predict and mitigate traffic jams caused by construction and other related projects. The pilot project would gather six months' worth of data from GPS tracking devices and service providers such as Uber and Careem to develop models to understand the different impacts of road disruptions, at different times during the day and in different parts of the city. As Lahore is governed by multiple authorities, and the data would be supplied by private operators, this initiative requires a significant amount of coordination and legal wrangling, which is currently underway.

Challenges to Using Data Effectively in Planning

This section provides a discussion of challenges to the use of data in urban planning in Pakistan. Based on interview responses, three sets of challenges are identified and described: data access and reliability challenges, data and technological literacy challenges, and institutional challenges. The role of e-government and policy is also discussed, with a brief overview of IT and data policies. The final section describes how urban planning itself is still limited in Pakistan, especially relative to Pakistan's rapid urbanization.

Data access and reliability challenges

Data availability

Nearly all interviewees spoke about the absence of data, especially for urban planning purposes. There is a severe lack of the right kind of data, which is reliable, organized, collected using well-designed instruments, available at the right scales, and in the machine-readable formats. An urban planner in Karachi mentioned that there was very little data on densification in the city, and that planners had to rely on crude estimates. His team is occasionally able to find sample data for specific projects, but that data is usually not extracted from a comprehensive data collection exercise and thus has limited validity.

Planners also found it difficult to get urban economic data; while national and provincial gross domestic product (GDP) was available, there was no GDP data at the district-level. Industrial and agricultural data could be used to make estimates, but service sector data was missing, despite contacting the State Bank of Pakistan and Federal Board of Revenue (which oversees tax collection).

Very little data is available at the scale of the city, which is the appropriate geographic scale for urban analysis; the best that can be provided is district-level data. Data is also collected at irregular intervals, limiting the quality of trend analysis. To take the same example, MICS survey data is available for 2004, 2008, 2011, 2014, and 2017 (Punjab Bureau of Statistics, 2018b).

Some data is simply not collected, such as housing needs; in other cases, data is not preserved. Hull (2012) reports being surprised that Islamabad, Pakistan's most carefully planned and modern city, had almost no records:

"The official in charge of CDA employee housing had no comprehensive documentation on how many housing units were under CDA control and where they were, though he managed perhaps as many as twenty thousand. A former CDA chairman told me that "there is no one who can tell you what [the] CDA owns. [P]ieces of land were acquired years ago and no one even knows we have them."" (p. 3)

(CDA is the Capital Development Authority, the corporation in charge of Islamabad's planning and development.)

Only one interviewee claimed there is sufficient reliable data, a senior bureaucrat in the Planning Department. He claimed that there was enough historical data to make reasonable estimates and implied that it was not the existence of data that mattered, but the degree to which data was used to inform policy-making.

Data access

"There is a lot of data in Pakistan, on papers and sitting in cabinets. It's there. Someone needs to pick it up, digitize it and form a single view of it. And then work on it. And that entire infrastructure investment of forming that single view is not there. It's a daunting task." (Interview, May 12, 2018)

Once the data exists, the second challenge is getting it. Accessibility is limited and haphazard as a lot of data is not digitized, and then is not shared.

Much of the existing data is inaccessible because it has never been digitized or made available in a machine-readable format. An Urban Unit manager told me about trying to get data regarding water quality a few years ago. The relevant department gave him four printed volumes because no electronic records were available. Another interviewee spoke about having to digitize 30 years of data for a single agriculture project.

A particularly egregious issue is data that is collected by organizations on behalf of government, but never shared on government websites. Consulting agencies, international development contractors, multilateral and bilateral funders, and a host of other actors routinely carry out data collection exercises in partnership with or on behalf of the Pakistani government, but that data is rarely accessible to anyone.

An example is the data that was collected for the Lahore Urban Transport Master Plan between 2010 to 2012, a project that cost USD $2.5 million (ALMEC, 2018; Alam, 2011). The Punjab government and Japan

International Cooperation Agency (JICA) contracted ALMEC Corporation and Oriental Consultants to create an urban transport master plan, along with an action plan and capacity development for implementation (JICA, 2012). This entailed a person trip survey and eleven different transport surveys, used to develop a transport demand analysis model with three different scenarios and forecasts. Satellite imagery was also used to generate new maps with layers indicating the road network, built up areas, and a number of other land use attributes. This data was placed in a transport database and a separate GIS database (ibid). Collectively, this exercise represents an unprecedented breadth and depth of updated and reliable information for traffic planning in Pakistan.

How was this data shared or used? One interviewee was a manager at the Transport Planning Unit at the time. He asserts that most of the data was effectively used for the urban transport master plan, and that it was shared with government. Unfortunately, the data was not shared in its entirety, nor in the right format; this was attributed to low department capacity and consultant malpractice (Interview, June 1). Today, neither the data nor the master plan is available on any government website; they can only be downloaded from the JICA website itself as PDFs.

This demonstrates another interviewee's claim, that only data that has already been compiled and is considered 'fit for reporting' is available online. Getting access to 'raw data' and micro-data is considered almost impossible without the help of insiders. Knowing where to find data is a skill in and of itself.

This inaccessibility is also due to an unwillingness to share data, which a number of interviewees highlighted. For a planner at the Urban Unit, this was the primary issue, and she spoke of a culture where sharing data was rare.

Data reliability

"Our research would be meaningless if we had to rely on official statistics." (Interview, May 9, 2018)

Even when data is collected and shared, it is frequently faulty. The data is often incorrect due to poor data collection practices. Data is often messy and needs to be cleaned, for which there are limited resources.

Standardized definitions are not used for indicators, which makes it difficult to interpret and compare datasets from different sources. For example, the definition of "urban" used for the PSLM survey is different from the definition of "urban" used for the national census (Pakistan Bureau of Statistics, 2018a).

Information is also often presented or used selectively, in keeping with a broad recognition that data is political, as identified in the last chapter. Major data collection exercises such as the census have been marked by political interference (Zaman, 2017; Interview, 2018). As a result, data is often viewed with skepticism and suspicion.

Due to the unreliability of existing data, planners must often resort to carrying out primary data collection on their own. This can be an expensive undertaking and take up significant time.

The unreliability and high costs of data limit its use in two ways: by directly limiting the use of data, and by creating a culture where data is not valued. For data to be used regularly, it has to be a trustworthy guide. In Pakistan, information can create more questions than it answers.

Ivan Sigal's reporting on the Karachi Circular Railway (KCR) highlights the confusing nature of information. There are five different data points for the length of the (KCR), four data points for the population of the city, and six data points for the number of buses in the city. Which number is correct? No one really knows.

"Karachi has 9,000 buses. It has 12,399 buses, but only 9,527 are running. It has 6,457 buses and 2,715 contract carriages, buses, and luxury coaches. There are 1,800 contract buses and 1,800 route buses. A few years ago it had 22,313 buses." (Sigal, 2018)

Data and technological literacy challenges

Data literacy was a theme that emerged early on in my research, and remained a prominent talking point throughout the interview process. Many interviewees talked about the lack of skilled human resources that fully understood the value of data, and could collect, store, analyze, and derive value from it. This was true for urban planners as well as the broader public service and was attributed to a lack of appropriate education and training. Technological literacy also emerged as a parallel theme over the course of interviews.

Data literacy among planners and in the broader public service

In general, urban planners have greater data literacy than public servants in general, due to the nature of their work. They undergo specific training in conducting surveys, analyzing data, and presenting it. These skills seem to be honed over the course of their careers. In practice, however, interviewees reported that junior urban planners and recent graduates had limited skillsets. In one instance, an urban planning agency hired a number of junior research assistants as interns and tasked them with extracting and consolidating data from different surveys. These research assistants made a number of basic errors, were unfamiliar with Microsoft Excel beyond a few basic functions, and forced senior research analysts to train them.

One of my interviewees was the head of a university department in architecture and planning. He mentioned that planners received training in data analysis, but were warned of possible shortcomings and there was a conscious effort to build their capacity to overcome the challenge of not getting the data.

Most interviewees agreed that at present, data literacy is limited at all levels of the broader public service. This is backed up by research; the Building Capacity to Use Research Evidence program surveyed over 1,500 public servants in Pakistan and found that most of them could not interpret a simple 2x2 table correctly (Callen et al., 2017).

The barriers to this data literacy are a lack of awareness and teaching capacity, vested interests and resistance to change, and the general low level of talent in the public service. Skilled talent can make much more money and exercise much more autonomy in the private sector, so it is difficult for the public sector to attract the people with data skills.

It should be noted that the discussion above does not refer to employees of the Pakistan Bureau of Statistics (PBS) or other government agencies specializing in data and statistics. According to an interviewee who has worked closely with the PBS, their staff are adequately competent and are hamstrung by political interference and institutional barriers more than anything else. In addition, there are certainly organizations, such as the Urban Unit, which boast a relatively high level of data literacy. Nearly everyone works with data in some fashion, especially research analysts and the GIS teams, and over time, staff develop a good feel for data and can identify which data is useful and which is redundant.

The experience of the Urban Unit's GIS teams is instructive. They developed country-leading expertise in geospatial technology in response to growing company needs, not as the outcome of a specific plan. The data collection teams used to do surveys and capture data manually, then started using Android smartphones in 2011. They developed a custom mobile application that could only share limited information. Over time the app was refined to add the ability to send pictures and GPS coordinates, and provide offline functionality. This required developing a unique emerging skillset, that of spatial information technology. According to the team lead, most IT specialists lack the ability to think in spatial terms, and GIS specialists lack the ability to do software development.

When the GIS teams began working, they faced two challenges: their own capacity was limited, and the technology was not ready either. Over the years, technology has progressed rapidly - more disk space, more powerful and faster processors, and advanced software. This has helped the team innovate in response to challenges, and simultaneously develop both data and technological capacity.

Lack of training and education regarding data

Low data literacy can be traced back to the poor quality of education, especially regarding statistics, which is at the core of data science. In business schools, statistics are taught primarily for market research purposes; in engineering schools, there is much more focus on applied mathematics. According to one interviewee, statistics is considered to be a joke by many engineering students. Another interviewee mentioned that computer science education in Pakistan is also largely outdated, with syllabi from decades ago still in use. An assessment of planning education programs also found limited training in statistics and information technology for planners (Ahmed, 2018).

In the public service, data literacy varies based on educational background. In the Central Superior Services exam, which is used to qualify and induct public servants into elite positions, statistics is an optional subject. For current public servants, IT modules are offered as part of training courses such as those offered by the National School of Public Policy. These modules are usually limited to using Microsoft Excel and PowerPoint, but the content has begun to change. Since 2014, the Punjab Information Technology Board has conducted a number of workshops on

data-driven decision making across the country, as part of the Mid Career Management Course (Information Technology University, 2021).

Technology literacy

Due to the role of technology in data collection and processing, technological literacy is a crucial element of data literacy.

According to a senior policy specialist, few public servants have adequate technical skills. In most government departments, technological literacy is usually limited to IT teams that are responsible for managing infrastructure such as computers, internet access, and servers. Organizations such as the Urban Unit and the Punjab Information Technology Board are rare and operate at arm's length.

Unfamiliarity or discomfort with technology means that tools to promote the use of data, such as decision support systems, may never be properly used. Urban Unit employees who had designed, produced and distributed these systems repeatedly stated that building the system is much easier than building the capacity to use the system. In other words, the real magic lies in change management, not in technological innovation.

An important element of successful technology adoption is securing buy-in from the top. But this is not enough. Hull describes how the CDA in Islamabad purchased an expensive computer system in 1996 to manage a database of land and compensation records. The effort to use electronic databases was championed by multiple chairmen of the CDA for years, but repeatedly failed as lower tiers of the organization refused to change their ways. No one would submit data to enter into the system installed in 1996 (Hull, 2012).

According to a senior manager at the Urban Unit, the system only changes when its processes are changed. The success of an IT intervention relies on making existing business processes dependent on the use of that IT system. If training and management pressure don't work, another tactic can be to change the laws. For example, there is no requirement to carry out spatial analysis before proposing a new school or hospital, unlike the well-established requirement to carry out environmental impact assessments. The manager told me that the Unit established GIS cells in twelve districts in Punjab and Khyber-Pakhtunkhwa to facilitate spatial analysis for

planning, but the technology was meaningless without a legal mechanism and enforcement.

Another barrier is that the public service can have unrealistic expectations of the technology. Without having a clear sense of their challenges and true requirements, they can request a technical fix, but the quality of the solution depends to a great deal on the quality of the problem definition and scoping.

Institutional challenges

The challenges of data collection, reliability, accessibility, and literacy are underpinned by institutional challenges. Efforts to use data are often foiled by vested interests in the status quo, distrust of technology, and the absence of a supportive environment and enabling policy.

Vested interests and distrust

In the anecdote about the failure to institute electronic databases in Islamabad, the point is not simply that top-down approaches don't work. The far more interesting point is that the effort to digitize land records destabilized existing power structures that relied on the old system.

Vested interests in the status quo block the use of data. Senior public servants who have been working on a particular file for a long time may believe that their accumulated wisdom supersedes what the data shows. They may also feel that their efforts are undermined or threatened by the introduction of new knowledge that could reflect poorly on their past work.

In addition, the use of big data and advanced technologies has the potential of displacing jobs in two ways: by identifying unproductive areas of work, and by directly automating tasks such as data analysis.

Another issue, especially regarding the use of big data, is trust. The application of big data requires technology and infrastructure investments; in a financially stressed country like Pakistan, there is little appetite to take risks in investing in something new. If positive outcomes are not immediately forthcoming, the investment is quickly deemed wasteful.

An unsupportive environment

Even when good data may be available and there is both capacity and willingness to use it, other elements of the system are often not ready for this change. This takes the form of both physical infrastructure as well as

existing processes. The use of data requires major reforms in terms of introducing and institutionalized evidence-based policymaking and e-governance, enabled by a supportive policy framework.

In terms of physical infrastructure, many public servants still do not have access to computers or smartphones, let alone questions of whether the computing hardware and software meet the needs or not.

Poor infrastructure limits data collection, as instruments for collection rely on the existence of certain conditions that enable these instruments to be used effectively. For example, the automatic number plate recognition system introduced by the Punjab Safe City Authority cannot read old or dirty license plates. This is such a challenge that the Authority has asked the Punjab government to change license plates to enable the e-challan system to work properly (Ahmed, 2017). This is by no means unique to Pakistan and echoes a challenge that Tesla and Volvo, two of the world's leading autonomous vehicle manufacturing companies, have faced. Their semi-autonomous vehicles relying on cameras and other sensors cannot navigate roads that have faded lane markings, even though two-thirds of roads in America are in poor condition (Sage, 2016). This highlights that the actual costs of introducing new technology are potentially quite higher than originally thought (Veoni, 2017).

Data access and reliability is also affected by government policy and how government departments that collect data are structured. For much of Pakistan's history, there have been a number of different agencies collecting data, at both the provincial and federal levels. The multiplicity of agencies can make it difficult to monitor data collection efforts, establish standards, ensure quality, and share data effectively. An example is spatial data in the federal government: since 2010, the federal government has been attempting to create a National Spatial Data Infrastructure (NSDI) (Ali & Imran, 2021). Despite the creation of a legal framework, the Surveying and Mapping Act (Gazette of Pakistan, 2014), as well as broad acceptance of the need for coordinated geospatial data management, the NSDI has not been able to be implemented.

In fact, the Surveying and Mapping Act and subsequent Rules passed in 2015 (Gazette of Pakistan, 2015) restrict the use of data. Clause 7 (Registration) of the Surveying and Mapping Rules reads as follows:

No public or private organization, private firm or individual, national or international, shall undertake any geospatial data collection, production or analysis work and surveying and mapping activities unless it is registered with the Survey of Pakistan.

Urban planning in Pakistan

Investigating the use of data in urban planning necessitates tackling both the use of data as well as urban planning in Pakistan. One of the most prominent research findings was that urban planning in Pakistan is itself a very limited endeavor. Most urbanization in the country has been unplanned, partly due to a lack of enabling policy for spatial land use planning. There are not enough trained planners, due in large part to the lack of career growth options. In addition, there is no culture of civic participation with planning and policymaking.

Much of the urban growth that has taken place has not been planned. Multiple urban development institutions have emerged over the years but remain mostly ineffective and helpless, due in large part to 'legal and institutional perplexities' (Ahmad & Anjum, 2011): lack of comprehensive, effective and updated legislation; overlapping, confusing and disempowered local governance; weak land markets; and limited capacity for execution (The Urban Unit, 2007; Yuen & Choi, 2012).

While discourse around the failures of urban policy tends to revolve around municipal delivery, such as road repairs, waste management, and water supply, spatial planning has been neglected as both cause and potential solution. Spatial planning in Pakistan is a patchwork of master plans developed for major cities, often relying on outdated assumptions by consultants parachuted in at the behest of international donors (Hasan, 2012). None of these plans have been implemented in full, and many were consigned to filing cabinets almost as soon as they were developed.

One of my most important and unexpected research findings was learning about exasperation with professional planning in Pakistan (Haque, 2018). Unprompted, several interviewees said that 'planning is dead here' and went on to speak about planning as an outmoded, inflexible field of professional practice with gatekeepers and limited opportunities for career growth and professional development.

Gatekeeping also limits the scope of planning practice; the longstanding dominance of architects means that their perspectives prevail in planning circles, informed by a different and sometimes narrower scope of expertise and knowledge than town planners. This can lead to conflicts in the way data is collected, analyzed, and interpreted. A review of master plans in Punjab reveals this influence, with plans criticized as being "essentially local, physical and restrictive" (The Urban Unit, 2007). The reliance on master plans, a mostly outdated approach, is itself emblematic of an outdated planning system.

Conclusion

While data is highly valued and recognized as essential in the planning process, it is poorly used, when it is used at all. The reality is that not a lot of data exists, and available data is highly questionable. In addition, it is not clear that existing data is being leveraged effectively in any case. Finally, data literacy among both planners and policymakers is extremely limited. What needs to be done, then, is to ensure that sufficient reliable data is collected and that there is capacity to leverage this data appropriately. This entails building the capacity of a number of organizations and, over the long-term creating a culture of data-driven decision-making. While there are a number of promising initiatives under way, they are not connected as part of a broader strategy and thus their effectiveness is likely to be limited. The biggest challenge to the use of data may be that urban planning in Pakistan is itself deficient in many ways.

References

Ahmad, N., & Anjum, G. A. (2012). Legal and institutional perplexities hampering the implementation of urban development plans in Pakistan. *Cities*, *29*(4),271-277. https://doi. org/10. 1016/j. cities.2011. 07.006

Ahmed, N. (2012). An evaluation of planning education in Karachi, Pakistan: Learning from the past to tread into the future. *Global Built Environment Review*, *7*(3), 1-21.

Ahmed, S. I. (2018). Mapping growth. *The News on Sunday*. http://tns.thenews.com.pk/mapping-growth/

Ahmed, Z. (2017, October 7). E-Challan system gets operational in Lahore. *Pakwired*. https://pakwired.com/e-challan-system-gets-operational-in-lahore/

Alam, A. (2011, March 11). An urban planning disaster in Lahore. *The Express Tribune*. https://tribune.com.pk/ story/131228/an-urban-planning-disaster-in-lahore/

Ali, A., & Imran, M. (2020). The Evolution of national spatial data infrastructure in Pakistan-implementation challenges and the way forward. *International Journal of Spatial Data Infrastructures Research*, *15*, 110-142.

ALMEC Corporation. (2018). *Project summary*. http://almec. org/ projects/ D04.html

Anwar, N. H. (2012). Thinking beyond 'engines of growth': Re-Conceptualizing urban planning discourse in Pakistan. *International Journal of Research in Architecture and Planning*, *15*(2), 223–238.

Batty, M. (1994). A chronicle of scientific planning: the Anglo-American modeling experience. *Journal of the American Planning Association*, *60*(1), 7-16. https://doi.org/10.1080/01944369408975546

Batty, M. (2013). *The new science of cities*. Cambridge, MIT Press.

Burton, I. (1963). The quantitative revolution and theoretical geography 1. *Canadian Geographer/Le Géographe Canadien*, *7*(4), 151-162. https://doi.org/10.1111/j.1541-0064.1963.tb00796.x

Callen, M., Khan, A., Khwaja, A., Liaqat, A., & Myers, E. (2017). These 3 barriers make it hard for policymakers to use the evidence that development researchers produce. *Washington Post*. https:// www.washingtonpost.com /news/monkey-cage/wp/2017/08/13/these-3-barriers-make-it-hard-for-policymakers-to-use-the-evidence-that-development-researchers-produce

Davoudi, S. (2006). Evidence-based planning: rhetoric and reality. *disP-The Planning Review*, *42*(165), 14-24. https://doi. org/10. 1080/ 02513625.2006.10556951

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*, *65*(3), 122–135. https://doi.org/10.1108/LR-06-2015-0061

Faludi, A., & Waterhout, B. (2006). Introducing evidence-based planning. DisP - *The Planning Review*, *42*(165), 4–13. https:// doi. org/10.1080/02513625.2006.10556950

Gazette of Pakistan. (2014). *Surveying and mapping act* (Pak.). https://na.gov.pk/uploads/documents/1402988824_562.pdf

Gazette of Pakistan. (2015). *Surveying and mapping rules* (Pak.). http://surveyofpakistan.gov.pk/SiteImage/Downloads/surveying_and_mapping_rules_2015.pdf

Hall, P. (2014). *Cities of tomorrow: An intellectual history of urban planning and design since 1880*. John Wiley & Sons.

Haque, N. (2018). *Urban planning from a planner's viewpoint (URDU)*. PodBean. https:// nhaquepod.podbean.com/e/urban-planning-from-a-planner-viewpoint/

Hasan, A. (2012, August 13). *Planning and Its assumptions* [Presentation]. http://arifhasan.org/presentations/planning-and-its-assumptions

Hull, M. (2012). Government of paper: The materiality of bureaucracy in urban Pakistan. University of California Press

Information Technology University. (2021). *Data driven decision making*. ITU. https://itu.edu.pk/it-capacity-building/data-driven-decision-making /

JICA. (2012). Lahore urban transport master plan, final report. *JICA*. https://openjicareport.jica.go.jp/pdf/12068110_01.pdf

Kwan, M. P., & Schwanen, T. (2009). Quantitative revolution 2: The critical (re) turn. *The Professional Geographer*, *61*(3), 283-291. https://doi.org/ 10.1080/00330120902931903

Pakistan Bureau of Statistics. (2018a). Methodology. *PBS*. http:// www.pbs.gov.pk/content/methodology-1

Punjab Bureau of Statistics. (2018b). Multiple indicator cluster survey (MICS). *PBS*. http://www.bos.gop.pk/mics

Sachs, W. (2010). *The development dictionary* (2nd ed.). Zed Books.

Sage, A. (2016, March 31). Where's the lane? Self-driving cars confused by shabby U.S. roadways. *Reuters*. https://www.reuters.com/article/us-autos-autonomous-infrastructure- insig/wheres-the-lane-self-driving-cars-confused-by-shabby-u-s-roadways-idUSKCN0WX131

Sigal, I. (2018). These studies led to further studies. *Places Journal*. https://doi.org/10.22269/180501

Simon, D. (2015, May 18). Mr. O'Malley's bad math. *The Audacity of Despair*. https://davidsimon.com/omalley-bad-math/

Thakuriah, P., Tilahun, N., & Zellner, M. (2017). *Seeing cities through big data*. Springer.

The Express Tribune. (2015, July 14). Fort Munro hill station set to become tourist haven. *The Express Tribune*. https:// tribune. com.pk/ story/921058/fort-munro-development-authority-hill-station-set-to-become-tourist-haven/

The Urban Unit. (2007). Assessment of land development and management practices in five large cities of Punjab. *Academia*. https:// www. academia.edu/11646173/Assessment_of_Land_Development_and_Management_Practices_in_Five_Large_Cities_of_Punjab

The Urban Unit. (2018). Punjab cities growth atlas 1995-2015. *Urban Unit*. http://uu.urbanunit.gov.pk/enewsletters/urbanatlas/index.html

Townsend, A. (2015). Cities of data: Examining the new urban science. *Public Culture*, *27*(2), 201-212. https:// doi. org/10. 1215/08992363-2841808

Veoni, D. (2017). Self-driving cars are coming, but US roads aren't ready for the change. *The Hill*. http://thehill.com/opinion/technology/353034-self-driving-cars-are-coming-but-us-roads-arent-ready-for-the-change

Wegener, M. (1994). Operational urban models state of the art. *Journal of the American planning Association*, *60*(1), 17-29. https:// doi. org/ 10.1080/01944369408975547

Yuen, B., & Choi, S. (2012). *Making spatial change in Pakistan cities growth enhancing*. World Bank. https: // openknowledge.worldbank. org/bitstream/handle/10986/17879/862470NWP0Worl0es0Growth0Enhancing.pdf?sequence=1&isAllowed=y

Zaman, F. (2018, September 19). Census 2017: How can flawed results have any credibility? *Dawn*. https://www.dawn.com/news/1358516

Zheng, Y., Capra, L., Wolfson, O., & Yang, H. (2014). Urban computing: Concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *5*(3), 1-55. https://doi.org/10.1145/2629592