Big Data and Urban Planning in Pakistan: A Case Study of The Urban Unit

by

Nabeel Shakeel Ahmed

supervised by

Dr. Murtaza Haider

A Major Paper

submitted to the Faculty of Environmental Studies

in partial fulfillment of the requirements for the degree of Master in Environmental Studies

York University, Toronto, Ontario, Canada

July 31, 2018

Table of Contents

Abstract

This Major Paper presents research on the use of big data for urban planning and smart cities in the global South. It does so through a case study of the use of data at the Urban Unit in the province of Punjab, Pakistan.

Based on a series of interviews and extended literature review, I trace the evolution of a science of cities and the growth of urban informatics and smart cities. I then define big data and discuss its related opportunities and limitations.

The bulk of this Paper consists of a case study of the Urban Unit and research findings regarding the use of data in planning in Pakistan. A number of challenges to the use of data are identified, classified into challenges regarding data access and reliability, data literacy, and institutional challenges. A major finding is that the practice of urban planning in Pakistan is quite limited in a number of ways.

The final chapter shares recommendations from interviewees and reflections on research findings, focusing on the politics of data. The paper ends by discussing future research directions.

Keywords: Big data, smart city, urban planning, evidence-based policy, Pakistan

Acknowledgements

I am grateful to many people for their support in helping me complete my Major Paper.

First and foremost, I am grateful to my advisor and academic mentor at FES, Dr. Abidin Kusno. Dr. Murtaza Haider, my supervisor, provided excellent guidance throughout the process and helped ensure my work was valuable. Dr. Ute Lehrer encouraged me throughout as well.

The Urban Unit team that I worked with last year, and who I interviewed for the paper, were all fantastic – especially Sana Riaz, Maryam Ibrahim, and Sarah Ahmad. Dr. Nasir Javed, the CEO of Urban Unit, provided critical support and helped me gather the necessary information.

All of my interviewees were generous with their time and knowledge. I came away from my fieldwork deeply impressed by the professionalism and passion of civil servants in Pakistan. In particular, I want to highlight the efforts and ingenuity of women working in the civil service.

A tiny village has enabled and cheered on my research. I will always be grateful to:

- Bianca Wylie, Melissa Goldstein, Howard Tam, and Andrew Do, for helping me understand how and why technology in the public domain matters.
- Asim Fayaz, for dismissing social enterprise as 'cute' and talking about TPI.
- Farzal Dojki, for reminding me to do something that will last a decade.
- Karim Harji, Jawad Jafry, and Kathy Bullock, for being personal and professional mentors.
- Loren March, Monica De Vera, Michael Romero, Aida Mas Baghaie, Alex Gatien, and Assya Moustaqim-Barrette, among many others, for being wonderful academic colleagues.
- Zainab Habib, Yousaf Shah, Ammara Niyaz, and Nabeel Hussain, for getting me into FES.
- Ayyaz Mallick and Shazma Abdulla for helping me get my bearings and stay radical.
- Chandrashri Pal, Maryam Khawaja, and Aamenah Ahmed, for helping me sort my notes.
- Urooj Shahzadi and Muriam Salman, for ensuring I focused and took care of myself.
- Reyna Ammar Khan and the vegetable patch, for being my daily mental health break.
- Ayesha Omer, Oni Ornan, Mariam Rafique, and Alyza Devraj for being there.

And most of all, my family for being my constant supporters and champions.

Foreword

Relation of Major Paper to Plan of Study

This paper is a part of my Plan of Study (POS) for completing the Master in Environmental Studies program. My Area of Concentration is 'big data and urban planning in the global South'; My POS examines the rise of the smart city paradigm and the use of technology in cities in the global South. In particular, I examine the use of big data by urban planners and policymakers as part of a broader drive towards modernizing urban planning and adopting evidence-based policy.

This paper aligns with all three components of my POS and helps me fulfill a number of Learning Objectives, which are as follows:

1. Urbanization and planning in the global South

Learning objective 1.1: To gain a solid understanding of the planning theories and practices of planners and policymakers in order to analyze government policy and programs by linking theory to practice.

Learning objective 1.3: To develop and build planning skills in order to better understand the approaches used by planners. I also wish to obtain the knowledge and skills necessary to meet the program requirements of the Ontario Professional Planners Institute for Candidate membership.

2. International development

Learning objective 2.1: To research the urban agenda in international development and aid work, focusing on NGOs and private consulting firms, in order to understand development trends.

Learning objective 2.2: Determine the role of public sector and community led initiatives in influencing urban development in order to better identify their roles and provide suggestions.

3. Big data and smart cities

Learning objective 3.1: Learn about the emergent ‘smart cities’ paradigm, especially in global South, in order to critically analyze the pros and cons of smart cities.

Learning objective 3.2: Explore the role and use of ‘big data’ in urban planning in order to understanding the potential and limitations of this approach.

Motivation

My interest in urban planning is a product of my experiences in Karachi, Pakistan, one of the largest megacities in the world. Karachi politicized me, turning me from a business student to a policy wonk. I left my career in the nonprofit sector because I felt the need to focus directly on the questions of urban planning, policy and politics. I chose to conduct my research on Pakistan because it was apparent that there was a glaring lack of attention on urban planning in one of the most highly populated and quickly urbanizing parts of the world.

I selected this topic because information technologies are an increasingly prominent component of planning practice. While skeptical of technological solutionism, I realized that the genie is out of the bottle; we will never go back to a world where data and smart cities are not part of the planning lexicon. My experience at the Urban Unit convinced me that something new and interesting was happening in the global South, but no one seemed to know about it.

My goal has been twofold: to tell the story of this phenomenon, and to make an original contribution and provide something of value to urban planners and policymakers in Pakistan.

Chapter 1: Introduction

Big data has become popular when discussing the future of cities. Indeed, both urbanization and technology seem to be both accelerating as increasingly dominant global forces. The confluence of these phenomena comes together most clearly in the notion of smart cities. Much of the debate around smart cities has revolved around the global North, that is to say, cities in the relatively prosperous and developed countries in Western Europe, North America, and Australia/New Zealand. The pace of urbanization is currently far more rapid in the global South, as economic growth and rural-urban migration fuel unprecedented shifts in the lives of hundreds of millions. Urban development in the global South has also witnessed the rapid uptake of new technologies. This transition is dynamic but has gone almost unnoticed; there are a number of gaps in our understanding of how technologies have been introduced, harnessed and implemented in the processes of urbanization that are unfolding across the global South.

This Major Paper addresses one of those gaps and investigates the use of big data in urban planning in the global South, using a case study of the province of Punjab, Pakistan. This chapter frames the paper by describing the key research questions and methodology, and by establishing the context of urbanization in Pakistan. Chapter 2 explores the role of technology in urban planning and introduces the concepts of big data and smart cities, providing both historical as well as local context. Chapter 3 presents the case study of the Urban Unit and shares research findings regarding the use of data in planning in Pakistan. Chapter 4 examines challenges to the use of data and the limitations of urban planning in Pakistan. Chapter 5 concludes the Major Paper by sharing recommendations from interviewees and reflections on research findings. The paper ends by discussing future research directions.

A. Research Questions and Methodology

I. Research questions

The primary research question of this paper is: *What are the opportunities and limitations of the use of big data by urban planners and policymakers in Pakistan?*

To answer this question, the following secondary questions have been identified:

1. What are the opportunities and limitations of big data in urban planning that have generally been identified?
2. What is the context of the introduction of big data and smart cities as dominant ideas in government?
3. How has data been used in urban contexts in Pakistan, and how could big data most effectively be implemented?
4. Could the use of big data transform urban planning in Pakistan, and if so, how?

One of the goals of this paper is to understand big data and smart cities, unpacking jargon and layers of technological utopianism by placing these terms in the appropriate context. These are chaotic concepts that are understood and hence used in different ways by an array of actors, including but not limited to the public service, politicians, private sector technology vendors, and academics. These actors often have competing interests but may choose to collaborate for a specific initiative, which risks long-term stability and success. It is necessary to understand the different incentives driving the promotion, adoption, execution and evaluation of big data and smart city initiatives.

Another goal of the paper is to situate these practices in the global South, especially in Pakistan. The type, volume and value of data that can be collected from Lahore's bus rapid transit system is very different from the data collected by the Toronto Transit Commission's network of subways, streetcars and buses. As is described in Chapter 2, big data has also been defined in relative terms, based on the contextual capacity for data analysis; thus, what is big data in Pakistan may not count as big data in Canada. Data and smart city as concepts may also emerge differently in the global South, with new actors such as international development agencies, funders and consultants. These concepts are thus laden with underlying assumptions that need to be identified, such as modernization and development.

While the prevailing narrative of 'developing countries' is that they are technologically backward and identified by a 'geography of lack', exploratory research in 2017 revealed that the global South is quite eager to harness the latest technologies to their fullest potential. There is a hunger to shed the labels of the past and leapfrog the long period of development that the West witnessed, often using extractive practices. As such, an important subset of research questions is about practice – identifying the ways in which big data transforms the practice of planning in Pakistan, with implications for the global South more generally.

Finally, a crosscutting theme of this paper is understanding the politics and governance of big data and smart cities, and the shift towards them. How is data produced and what is its value? Who owns and controls data, and for what ends? What are the implications for participation, decision-making, and democracy?

The paper begins with an extensive literature review tracing the genealogy of big data and smart cities. Much of the paper is about uncovering how data is used in planning in Punjab, focusing

on the Urban Unit, a provincial urban planning and consulting agency that has developed expertise in the use of data and technology. A number of examples of how the Urban Unit and other organizations are promoting the use of data and analysis will be highlighted, with three initiatives described in detail in the appendix. These examples shed light on the kinds of urban challenges that big data can help solve for Pakistani policymakers. The interview findings will be supplemented with commentary throughout, relating to the questions and themes identified above. The paper will conclude by identifying best practices for the use of big data in Pakistan, as well as future research directions.

II. Methodology

The research design was informed by exploratory research conducted during the summer of 2017, while working at the Urban Unit for four months. This helped identify key questions as well as potential case studies.

A mixed-methods approach is implemented in this research, combining a literature review and semi-structured interviews. 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. Five participants are current employees of the Urban Unit; two others are senior public servants with the Planning & Development Department and the Punjab Safe City Authority; the remaining two are an urban planner, and the CEO of 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.

I also reviewed a number of data sources that are used by planners and policymakers, and the publications they have produced. Data sources include the Pakistan Social and Living Standards Measurement Survey (PSLM, 2015), Labour Force Survey (LFS, 2015), and Multiple Indicator Cluster Survey (MICS, 2014). Urban Unit publications include the Urban Atlas, Urban Gazette, Urban Geographic, and Pakistan Journal of Urban Affairs. While not approaching the standard of a content analysis, this review helped conceptualize the paper and identify gaps between discourse and practice.

This mixed-methods approach, centered around a case study, has a number of limitations. Notably, findings are likely biased as a purposive sampling approach was implemented, and most of my interviewees represented either the Urban Unit itself or the government. I was not able to formally interview someone at the Punjab Information Technology Board (which has been a key actor in the Punjab government's adoption of technology over the last several years). The focus on a case study also limits the broader applicability of the research findings.

My own positionality undoubtedly affected both my chosen methods and analysis. My relationship with Pakistan, although deep and enduring, has largely been formed in absentia and limits my understanding of the Pakistani state in important ways. As someone relatively privileged in terms of socioeconomic class and gender, especially in Pakistan, it is easier for me to navigate corridors of power. I have a number of biases and blind spots, such as being prone to Western ideological constructs. This paper does not devote sustained attention, if at all, to the questions of class, gender and climate change, even as urbanization is salient for all three. These are important areas that merit further study.

B. Urbanization in Pakistan

"The Future is Urban", declared the United Nations in 1994, and indeed the world crossed a threshold sometime over the last decade and more than half of humanity lives in cities, in a first for humankind. Pakistan is on track to follow that lead and it is estimated that over 50% of its population will be urban by 2030 (United Nations, 1994; Kugelman, 2014).

Over the last decade, Pakistan's urbanization has received increasing attention from both local and international observers. It is the sixth-most populous country in the world, and the most urbanized nation in South Asia. Its dynamic urbanization implies major demographic and geographic shifts, with deep spatial, economic, social, and indeed cultural implications.

The 6th Population and Housing Census 2017 indicated that 36.7% of the population is urban, growing at a rate of 2.96% since the last census in 1998 (Pasha, 2018). As can be seen in the table below, this is the highest share of urbanization in the region. It is also noteworthy that Pakistan's population growth rate is the highest in the region by far.

Comparison of Population Indicators of Pakistan in South Asia				
	Population (2016)	**Growth rate**	**Urban share**	**Urban growth****
India	1,320m	1.43%	33.1%	2.6-2.4%
Pakistan (World Bank)	193m	2.11%	39.2%	3.3-3.2%
Pakistan (2017 Census)	208m	2.38%	36.4%	3.0%
Bangladesh	163m	1.43%	35.0%	3.8-3.2%
Nepal	29m	1.32%	19.0%	6.0-3.2%
Sri Lanka	21m	0.71%	18.4%	0.6-1.4%

Table 1. Extracted from Pasha (2018); Source: World Bank, World Development Indicators
***Decline in annual growth rate to 2016 growth rate*

As Pasha (2018) has identified, the census numbers may actually be understated. Government and World Bank estimates were both upwards of 3 percentage points higher than the census findings, in keeping with the urban growth rate observed in the previous census. In addition, there are concerns that the definition of urban used in the census may lead to distortion and undercounting, as it relies on the urban areas defined as administrative units by provincial governments. Specifically, this method does not account for urban development in the periphery (Anwar and Ahmed, 2017). The release of the provisional findings of the 2017 census generated intense debate about its methodology, and the findings were eventually rejected by the Sindh government (The News, 2018). The 2017 census thus reflects a broad set of concerns around the quality as well as the politics of data that are explored below.

Qadeer (2014) has also argued that relying on the census definition significantly misrepresents Pakistan's urban nature. He points to what he refers to as urbanization by implosion, the growth of rural populations to urban density levels. If we apply the population density criteria used by UN Habitat, Pakistan is already well over 50% urban on a purely spatial basis (ibid; Ali, 2013). It is equally important to understand Pakistan as an urbanized society, one where life and economy has been transformed by urbanization (Qadeer, 1999).

Irrespective of the debate regarding the numbers of urbanization, there is general consensus on the factors leading to urbanization, as well as some of the likely implications of this urbanization.

The first factor, as hinted above, is natural growth due to the prevailing high population growth rate and fertility rate, even though it is likely lower in urban areas (Martine et al, 2013). The second factor is rural-urban migration, most of which is intra-provincial in nature (Nabi, 2015; Kugelman, 2014; Jan et al, 2008). The two major reasons commonly identified for rural-urban

migration have been employment and economic mobility (Hasan & Raza, 2013; Siddiqui, 2014; Iqbal, 2014) and conflict (Haider, 2014). There is emerging evidence of the significance of climate change and related disasters (Saeed et al, 2016) and conflicts such as the global war on terror and battle against extremism (Hasan, 2012; Yousafzai, 2013; Dominguez, 2014).

Urbanization is not just a result of some of Pakistan's existing challenges, it also presents an opportunity to address some of these challenges. Around the world, urban areas are associated with greater prosperity, better livelihoods, and improved social outcomes as they concentrate employment opportunities and help government allocate resources efficiently. Urban development, especially when compact, can also address factors leading to climate change. For Ul Haque (2014), cities are the engines of growth that can revitalize Pakistan. Global institutions such as the UN and World Bank promote this point of view, trumpeting the role of cities in addressing national challenges in the New Urban Agenda (United Nations, 2017).

The urban transition has not delivered as promised, however, and may indeed be leading to greater challenges. For one, cities in Pakistan produce a comparatively lower level of national GDP than regional counterparts (Nabi, 2015). In addition, poor planning and service delivery has exacerbated issues such as housing, transportation, and water and sanitation (Kugelman, 2014).

Poor urban outcomes are not just the result of poor implementation, they are also linked to governance and political challenges. A major issue is the failure to develop, or devolve power to, institutions that can meet Pakistan's urban needs, especially at lower tiers of government (Ali & Rumi, 2013; Qadeer, 2014). What institutions exist are often centrally managed upwards rather than being responsive 'downwards', with limited citizen accountability (Ali & Rumi, 2013). These failures stem from a national policy agenda that ignores urban imperatives and the

opportunities of urban renewal (Qadeer, 2014; Ul Haque, 2014). This ignorance is not accidental, argue Haider and Badami (2010). In fact, local governments have failed precisely because they have been instrumentalized by federal governments, both military as well as civilian, in their struggle for power. As Pakistan reaches a decade of uninterrupted (albeit fragile) democracy, the recent elections in July 2018 marked the first time that two back to back governments completed their tenure. It remains to be seen how policymakers will treat the urban question in the future.

These challenges bring us back to focus on the potential of adopting data-driven approaches and moving towards smart cities. A key theme of this paper is that this adoption 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 & 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 paper 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. These risks are explored further below.

Chapter 2: Big Data and Urban Science

This chapter describes the role of big data in urban planning, placing it in the context of "the new science of cities" (Batty, 2013) and identifying the use of big data as a part of the smart city paradigm. The chapter begins by describing the evolution of a science of cities and urban informatics, also placing big data in the context of smart cities. This is followed by defining big data, identifying some notable applications in urban planning, as well as the opportunities and limitations of big data. The chapter concludes by sharing perspectives on big data from interviewees in Pakistan.

A. The New Science of Cities

The use of data and technology is not a new phenomenon in the history of cities or urban planning, but has undergone a number of changes over the years. Certainly, planning has always been linked to the application of scientific knowledge, seen as 'rational social action' (Escobar, p.148, cited in Sachs, 2010). Faludi & 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). A decisive shift occurred when urban development accelerated in the 1950s, resulting in what Hall called 'the systems revolution' (Hall, 2014). Cities were recognized as not just places and spaces but also as networks and flows. Locations were more than simply physical destinations, and understood as nodes of interactions. Crucially, in this systems paradigm, "cities are more like biological than mechanical systems" (Batty, 2013, p. 14).

Planning changed from being understood as an outgrowth of architecture to being recognized as an applied science in its own right (Klosterman, 1994).

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[1] (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, p. 394).

This period reached a turning point, especially in the United States, 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). Douglass Lee's influential 'Requiem for Large-Scale Models' (1973) pointed out a number of

[1] While bitter debates with critical geographers in the 1970s may have given the impression that quantitative approaches were done and dusted, there has been a 'critical (re)turn)' to quantitative geography, recognizing that techniques such as locational analysis and spatial science were developed with progressive motives and now embodied in critical GIS (Kwan & Schwanen, 2009).

underlying flaws in the urban modeling approach (such as excessive ambition, lack of fine grained data, and prohibitively high costs) and seemed to herald the end of an era.

As Batty (1994) pointed out, it was not so much the end of an era as much as a reality check. Modeling and other applications of science and technology continued to receive attention from working groups conducting research projects around the world, but the scale of this work shifted from that of transforming the city to improving its management and operations. 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).

It is not only the process of planning that has been affected; ICTs and the 'datafication' of the city (Mayer-Schoenberger & Cukier, 2013) have also transformed the nature of the urban itself:

> "...computers which were once thought of as solely being instruments for a better understanding, for science, are rapidly becoming part of the infrastructure [of the city] itself, controlling new infrastructure through their software, influencing the use of that infrastructure, and thus affecting space and location." (Batty, 1997, p. 159)

While Batty emphasizes the locational implications of ICT, Madera (2010) notes that personal computing has transformed citizens, citizenship, and the nature of connections and social organization. For Rabari and Storper, this is the ‘digital skin of cities’, which consists of a vast, enveloping network of sensors and communication technologies, and whose implications extend beyond the production of big data and ‘smart’ urban management to public participation, governance, and a digitally mediated society (2015). This is the digital city, the smart city, the sensored city, the metered city, the real-time city, and the intelligent city.

These developments have led to what Batty (2013) calls 'the new science of cities', with expanding academic literatures in urban operations and governance and the new ones of urban informatics (Foth, 2009), urban computing (Zheng, 2014), intelligent cities (Komninos, 2009), and smart cities (Townsend, 2013). Townsend summarized some of the shifts associated with this new science (in Foth, 2009):

- New opportunities for decentralized, *bottom-up* planning approaches, in contrast to the *top-down* approaches favored by earlier generations of planners. This echoes Batty's note that systems thinking inherently questions the top-down approach to urban planning (Batty, 2013, p. 26).
- A focus on the *software* as well as the *hardware* of cities; uncovering and dealing more directly with not just the physical land and built infrastructures of the urban, but also its information flows and processes
- The ability to glean fine-grained insights and apply tools selectively; using a *scalpel* instead of Robert Moses' *meat ax* (Caro, 1974, p. 849)
- Unprecedented *transdisciplinarity* - adding computer, information and data sciences to the already crowded realm of urban studies

It is especially important to place big data in the context of urban informatics and smart cities. Urban informatics has been situated at the intersection of place, technology and people and defined by Foth, Choi & Satchell as "the study, design, and practice of urban experiences across different urban contexts that are created by new opportunities of real-time, ubiquitous technology and the augmentation that mediates the physical and digital layers of people networks and urban infrastructures" (2011, p. 4). With origins in ubiquitous computing, it focuses on the use of information and application of technology in cities. An early compendium of work highlights technology-enabled participation in urban planning, location science and remote sensing, linking the physical and the digital urban, and wireless and mobile technologies (Foth, 2009).

Urban informatics is a broader approach than 'urban computing', which is more technical and is explicitly focused on solving problems (Zeng et al, 2014). A broader definition of urban informatics by Thakuriah et al (2017) mentions the need to integrate theoretical and empirical insights and acknowledge the political economy of urban informatics as well. This paper integrates applications of, and lessons from, both urban computing and urban informatics.

A wide range of applications have been identified and tested for the new science of cities. 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).

Big data and smart cities

Big data in the urban realm can also be conceptualized part of a broader shift to smart cities, which must be investigated and has developed its own rich vein of scholarship (see, for example, Mora et al, 2017).

At the broadest level, the idea of a smart city revolves around the use of technology to improve urban life; Townsend defines them as

> "places where information technology is combined with infrastructure, architecture, everyday objects and our own bodies to address social, economic and environmental problems." (2013, p. 15).

This deceptively simple definition hides two distinct approaches. Much of the talk of smart cities consists of a decidedly corporate techno-utopian intervention into the discourse around urban policy and planning (Hollands, 2008; Wiig, 2015; Mora et al, 2017; Future Cities Catapult, 2017). Hollands argued forcefully that the notion of a smart city was a powerful ideological construct, combining the use of networked infrastructure and ICTs with business-led urban

development (2008). This notion of 21st-century urban entrepreneurialism has serious implications for social polarization, leading to greater inequity and disparities. A different perspective is to understand the smart city as one where technology enables greater inclusion and participation, and where the exciting developments are not top-down transformations brought about by shiny computing investments, but the wisdom of the crowds being tapped to nurture numerous bottom-up interventions (Nam & Pardo, 2011; Kitchin, 2014a; Hollands, 2015; Future Cities Catapult, 2017). This perspective also risks falling into the trap of what Krivy refers to as 'cybernetic urbanism', where citizens act as sensing nodes for an environmental-behavioural model of control (2018).

While smart cities are a global phenomenon, their emergence is undertheorized in the context of the global South. There is virtually no published academic literature on the phenomenon of smart cities in Pakistan. Ayona Datta's research on smart cities in India comes closest, highlighting how postcolonial urbanization in South Asia emphasizes techno-modernity (2015).

In this context, the New Urban Agenda and United Nations Sustainable Development Goals also link the data revolution to development practice and urban planning in the global South. The urban sustainable development goal (USDG) emphasizes the use of data, measurement and metrics, highlighting a notion that there are elements of urbanism that can be standardized. While this approach faces a number of practical and political challenges around the collection and use of data for indicator-driven approaches, it also represents an opportunity to move past the failure of the Millennium Development Goals if notions of expertise and measurement are critically explored (Klopp & Petretta, 2017; Caprotti et al, 2017). In addition, the global consensus developed at UN Habitat III (the latest UN global summit to focus on sustainable urbanization) subtly undermines the measurement focus of the sustainable development goals. The Quito

Papers represented a shift from modernist Charter of Athens, moving from viewing cities as machines to seeing them as open systems, implying that the idea of data-driven, smart cities cannot simply be applied as a top-down maneuver but as a contextually aware and locally responsive set of tools (Trundle et al, 2016).

B. Big Data and its applications

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. This section provides an overview of big data and its applications, opportunities and limitations.

I. Defining big data

Big data has become very popular as a term and as a concept over the last decade, across academic, corporate and public-sector environments (including urban planning). We are in the early stages of what is commonly referred to as "the data revolution", and urban planners in particular are moving from a historically data-poor environment to a data-rich one (Kitchin, 2014 Data revolution). The speed with which big data has gained popularity has led to a multiplicity of approaches and definitions. This section defines big data and identifies big data applications in urban planning. This is followed by a brief discussion of the opportunities and challenges provided by big data.

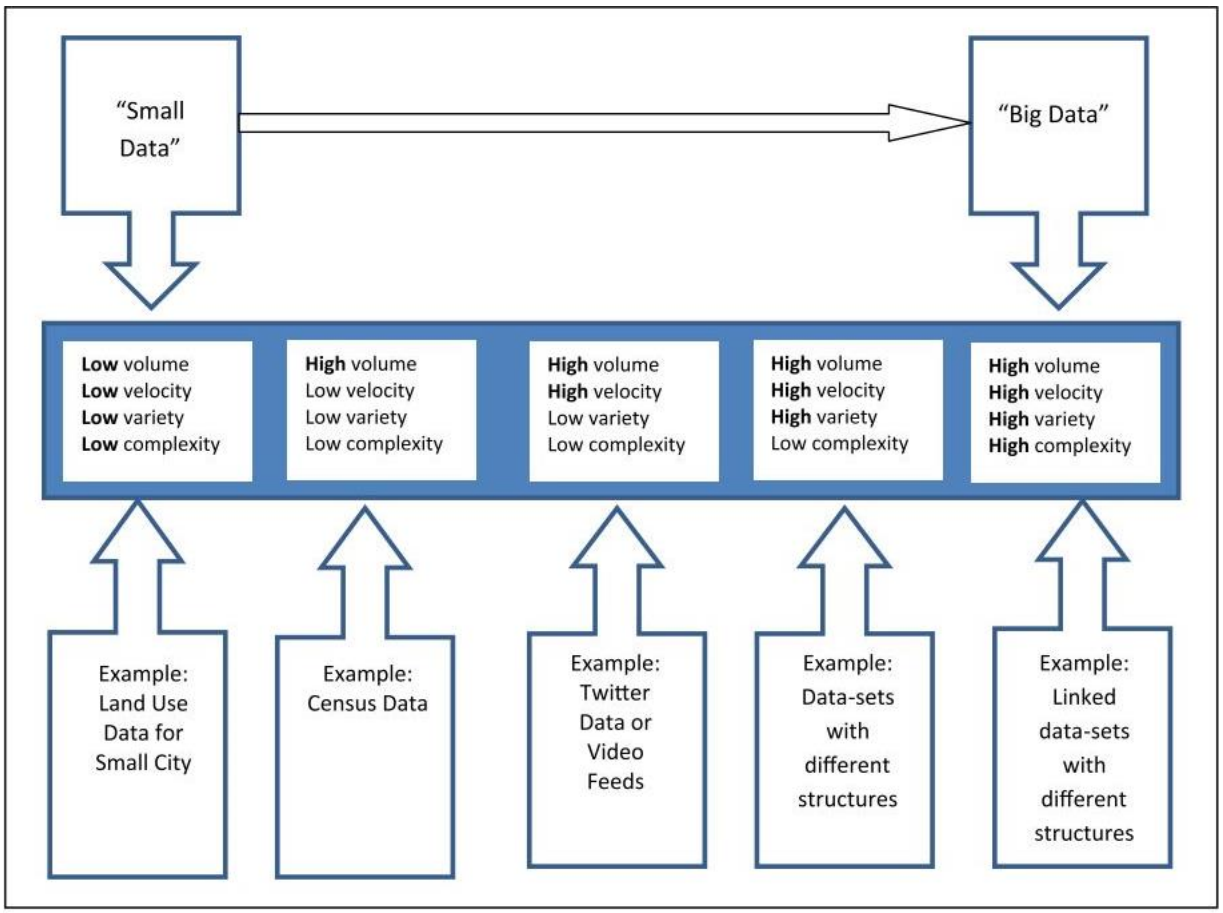


Figure 1. Data continuum (from DeSouza & Jacob, 2017).

As implied above, big data is not just a term, but also a phenomenon and a discipline (Diebold, 2012). As a phenomenon it reflects the shift in data that is captured, stored and used in a digital era where almost every interaction and event can be captured and recorded. The world's data volume has expanded exponentially over the last few decades; according to IBM (2016), 90% of the data in the world had been created in the previous two years alone. This growth is driven by a massive increase in the number of people using the internet - from 2.4 billion in 2014 to 3.8 billion in 2017 (Schultz, 2017) - as well as an explosion in the number of devices that capture data, described by the Internet of Things (Cisco, 2017). Data is now measured in zettabytes (a trillion gigabytes), with stored data volume set to almost triple from 1.8 ZB in 2016 to an

estimated 7.2 ZB in 2021 (Cisco, 2017). Significantly, much of this growth is coming from emerging markets, with the centre of gravity of data generation moving away from mature markets to emerging ones (IDC, 2014).

This 'tsunami' of data has led to the formal definition and discipline of Big Data. While a number of definitions exist, the most enduring description has come from Laney (2001) who defined the three Vs of big data: volume, variety, and velocity. A number of other dimensions of big data have been identified, but the point is that big data cannot simply be defined in terms of size, but must incorporate a number of characteristics. Volume is certainly at the heart of it, and refers to the sheer magnitude of data as outlined above. This aspect of the definition is relative, as it refers to the ability to process data. A personal computer would not have been able to handle 32 GB of data in 1990, but most modern smartphones can deal with this volume at ease. Variety refers to the nature of data; big data can be structured (like census data stored in a database), unstructured (such as text messages, photos and audio files), and semi-structured, which can be a range of data types that do not meet certain standards and usually have limited structural characteristics, such as sensor data that has associated meta-data. Velocity, the third major dimension, refers to the speed at which this new data is produced and has to be processed. Over 400 hours of videos are uploaded to YouTube on every minute of every day (Schultz, 2017); every tap of a smart card on a bus or subway produces another data point, thousands every second in large transit systems (Zikopoulos et al, 2013; Haider & Gandomi, 2015). For Kitchin & McArdle (2016), velocity is the key differentiator as it accounts for volume as well.

A number of additional dimensions have been identified, such as complexity and value. De Mauro et al (2016) conducted a literature review and found that four themes emerged when describing big data: information, technology, impact, and methods. They noted that big data

cannot be understood in the absence of specific methods and purpose, and proposed that Big Data should be 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." (2016, p. 131)

This attribute-based definition is the most popular, but by no means the only way Big Data is understood. It has been noted that big data is not only quantitatively different, but also qualitatively different from other types of data (Andreessen Horowitz, 2017). Kitchin & McArdle (2016) contend that there are multiple forms of Big Data and that it is important to distinguish these in order to make best use of it. The context makes all the difference.

- Software giant Cisco defined big data as "data deployed in a distributed processing and storage environment", noting that such an environment is used when certain 3 V thresholds are met (2017).
- In the science & technology studies field, boyd & Crawford treated Big Data as a phenomenon and defined it as "a cultural, technological, and scholarly phenomenon that rests on the interplay of technology, analysis, and mythology" (2012, p. 663).
- Thakuriah et al (2017) provide a broader, non-technical definition, calling it "structured and unstructured data generated naturally as a part of transactional, operational, planning and social activities, or the linkage of such data to purposefully designed data". (p. 14)

This paper relies on the definition provided by De Mauro et al, but as will be seen later on, the common understanding of most interviewees was much closer to the definition provided by Thakuriah et al.

II. Sources of big data

Big data are gathered from a number of different sources. Thakuriah et al (2017) classify the sources of urban big data based on typical user communities, which is a convenient method of classification. These sources include:

1. Infrastructure-based sensors, such as water usage, smart grids, and automotive sensors.
2. User-generated content and 'social' sensors, including social media, 'self-quantified' data such as that produced from wearable technology and smartphone trackers, and participatory sensing systems.
3. Government administrative data, including open data such as property tax records and transit records, and confidential micro-data such as social, health and education records.
4. Private sector data from both business-to-consumer as well as business-to-business perspectives, such as transaction records and customer profiles, and operational and management records such as factory production levels and stock trading data.
5. Arts and humanities data, such as images, music, other media and cultural products.
6. Hybrid data, such as census and survey data produced by government statistical bodies as well as private research firms.

A different approach is to consider the sources of big data based on ownership and access, which has important implications for governance and privacy in particular.

III. Big data management & analysis

As may be expected, not all of big data is useful; indeed, Cisco predicts that only 10% of the big data stored by 2021 will be useful (2017). An IBM book on big data makes the point that "Big

Data is all about better analytics...no one has ever delivered a single penny of value out of storing data" (Zikopoulos et al, 2013, p. 4).

There are a number of techniques that are used to extract value from data. Two major stages can be identified: data management and data analytics. Data management includes data acquisition, recording and warehousing; data mining, extraction, cleaning and annotation; and data integration, aggregation and representation (Haider & Gandomi, 2015; Sivarajah et al, 2017). Big data analytics (BDA) is a growing field in its own right and can be split into data analysis and modeling, and data interpretation.

Figure 2. Extracted from Haider & Gandomi, 2015; Sivarajah et al, 2017.

It is expected that BDA will grow in sophistication and, enabled by more powerful processing capabilities and technological innovation, lead to new techniques being developed, such as real-time analytics. Already there is a new generation of tools being used by big data practitioners, which includes data mining tools, batch-based processing tools, and stream-based processing tools. The table below identifies some of these tools. Machine learning and pattern recognition are emergent fields relying on big data that have received increasing attention in the last decade. Artificial intelligence and blockchain technology are also linked to or enabled by big data, but their use in urban informatics is presently limited. For reasons of space, these are not explicitly described in this paper.

Batch-based processing tools	Description
Hadoop	To perform the processing of data-intensive applications
Skytree Server	To process large amounts of data at high speed
Talend Open Studio	To provide a graphical environment to conduct an analysis for big data applications
Jaspersoft	To produce a report from database columns
Dryad	To improve the parallel and distributed programs and scale up the capability of processing from a small to a large number of nodes
Pentaho	To generate reports from a large volume of structured and unstructured data
Tableau	To process large amounts of datasets
Karmasphere	To perform business analysis
Stream-based processing tools	**Description**
Storm	To perform real-time processing of massive amounts of data
Splunk	To capture indexes and correlates real-time data with the aim of generating reports, alerts, and visualizations from the repositories
S4	To process unbounded data streams efficiently
SAP HANA	To provide real-time analysis of business processes
SQLstream s-Server	To analyze a large volume of services and log files data in real-time
Apache Kafka	To manage large amounts of streaming data through in-memory analytics for decision-making

Table 2. Extracted from Yaqoob et al (2016).

In addition, three kinds of analysis are identified by Sivarajah et al (2017):

1. *Descriptive analytics*, which help understand what happened. This relies on the use of statistical methods and business intelligence tools such as dashboards, scorecards, and data visualization. In urban planning, decision support systems are especially popular.

2. *Predictive analytics*, which help understand what is going to happen. Statistical methods for forecasting are applied here, such as regression techniques and machine learning techniques (Haider & Gandomi, 2015).
3. *Prescriptive analytics*, which help make recommendations for the future. These rely on establishing cause-effect relationships and evaluating the impact of simulations, for example by asking what-if questions.

IV. Big data applications

Given the broad scope of data that can be gathered and the diversity of analytical tools and techniques available at the disposal of urban planners and policymakers, there is a seemingly limitless range of applications for big data in cities.

Big data can provide new insights into longstanding problems, such as those of urban form, transport modal choices, the housing market, and labor mobility. It can also provide the ability to conduct new forms of empirical research and implement projects at new scales (Thakuriah et al, 2017). For example, planners can deploy sensors and develop a smart grid to accurately track the use of utilities such as electricity and water (Diamantoulakis et al, 2015).

Transportation planning is an area where big data can have a meaningful impact. Intelligent transportation systems can be upgraded and transformed into 'data-driven intelligent transportation systems' (Zhang et al, 2011) that can greatly enhance the efficiency of public transit systems, improve traffic flow, and improve public safety. Transport planners can leverage data collected from automated fare collection systems to infer passenger travel patterns, origins and destinations and conduct agent-based transport modelling (Tao et al, 2014; Anda et al, 2017). Planners can collect bike data to plan bike lanes that effectively account for real-world

constraints (Bao et al, 2017). An example of a new use of big data is predicting crowd flows for traffic management and public risk assessment (Hoang et al, 2016).

Planning support systems (PSS) and urban dashboards are specific tools utilizing big data that have received great interest over the years. They are briefly reviewed here as they are being implemented by the Urban Unit, which will be revisited in the following chapter.

PSS provide a framework to realize the dream of large-scale urban modeling, which initially attracted great excitement but ended up following a prosaic path, as described above. They are defined as geoinformation tools that combine the use of technology and computer-based tools with geospatial information specifically for planning purposes (Geertman & Stillwell, 2004; Vonk et al, 2005; Pelzer et al, 2014).

PSS are distinct from decision support systems, but incorporate several aspects of the latter; the examples in the case study include both. While PSS are dedicated to planning processes and tasks, DSS can be designed to support a wider range of decision research processes. While the former are used by professional planners for long-range problems, the latter are more suitable for executive decision-makers at shorter time-frames (Geertman & Stillwell, 2004).

One example of a PSS is *What If?*, a GIS-based online system that is currently used in 22 countries. It can project land use, housing and employment and can help ascertain the suitability of different projected uses (Klosterman, 2008; Pettit et al, 2013; Pettit et al, 2018; What If?, 2018). Another highly cited example is UrbanSim, a model for analyzing urban growth and transport-land use interactions (Waddell, 2002). UrbanSim is now used in over 70 countries and has its own open-source toolkit with tools such as ActivitySim, which simulates activity-based travel, and Spandex, which handles spatial analysis and data extraction (UrbanSim, 2018).

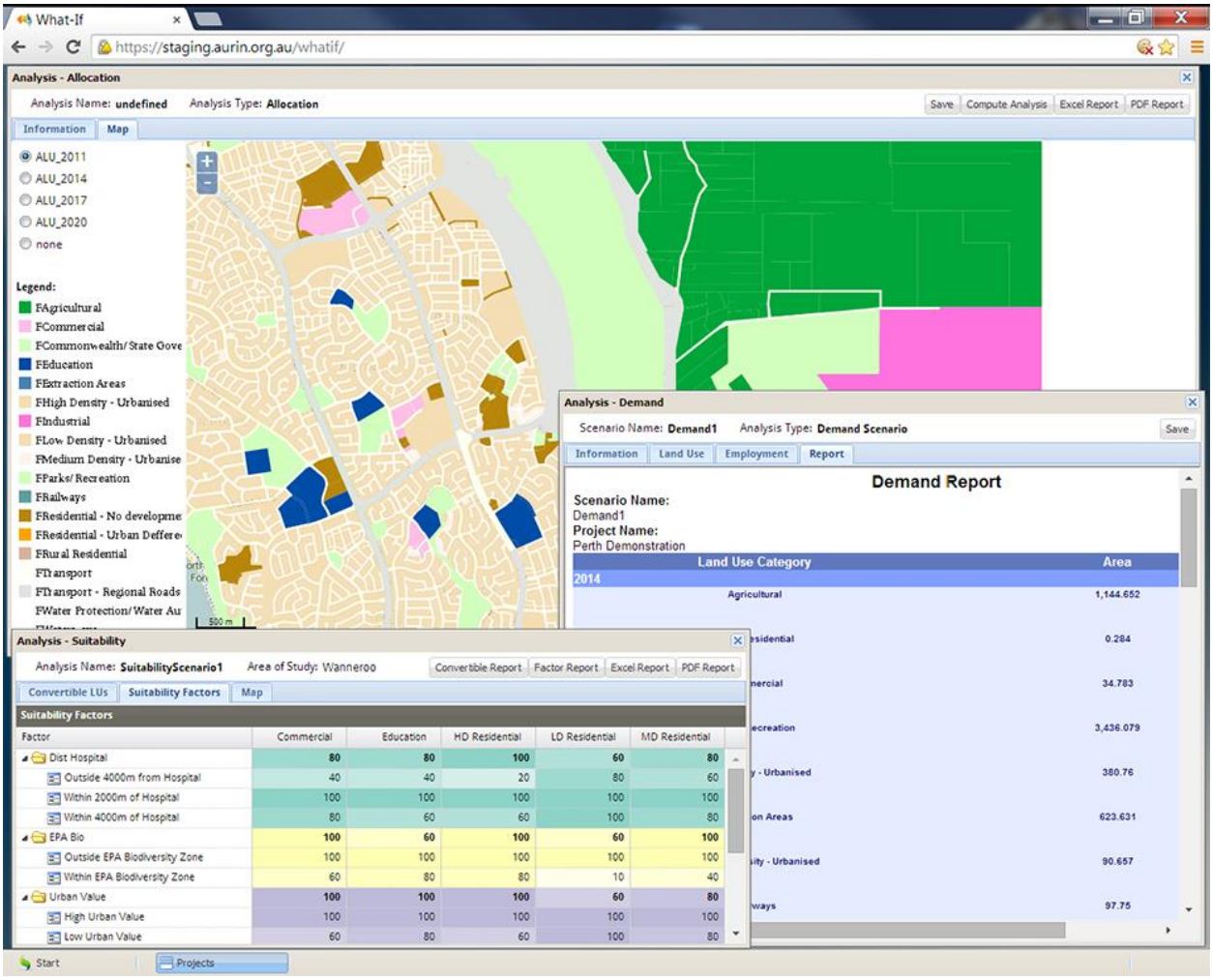


Figure 3. Screenshot of What If? (Roveda, 2017)

Despite the increasing sophistication and accessibility of PSS, they have not received widespread adoption. This 'implementation gap' (Geertman, 2017) has persisted for decades due to a number of factors, most significantly a lack of awareness among practitioners, perception of complexity (the 'black box' issue), and lack of appropriate training and experience regarding PSS (Vonk et al, 2005). Perhaps most significantly, the tools available were simply not ready - they did not meet the diverse needs of planners, especially land-use planners, and were not cheap and simple enough to use. This may change: the increased quality and quantity of PSS options, greater openness to technology in the public sector, perception of added value in terms of collaboration,

and advent of smart cities suggests potential for greater PSS uptake than before (Pelzer, 2014; Geertman, 2017; Pettit et al, 2018).

Dashboards are visual tools that consolidate and display information on a single interface. They can incorporate multiple different forms of information and multimedia, such as data tables, charts and graphs, maps, and live video feeds. These dashboards have dynamic capabilities, updating data in real-time and allowing users to filter data, apply queries, and finely customize their view (Kitchin et al, 2015; Gray et al, 2016).

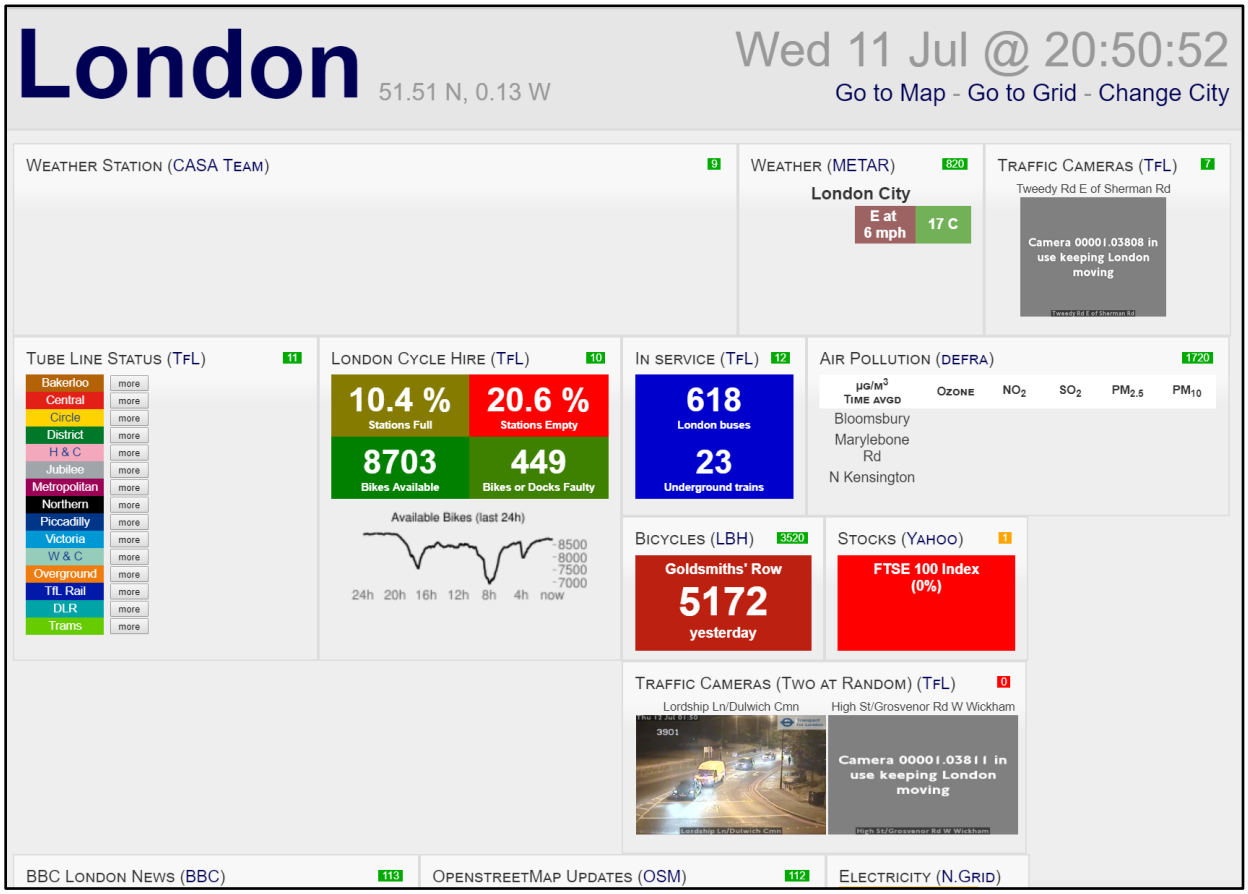


Figure 4. Screenshot of London dashboard (CityDashboard.org, 2018)

Dashboards have utility in a context where information is quantified and represented in the form of indicators, and where benchmarks are used to assess performance. Urban indicators represent a 'civic epistemology of quantification' (Miller, 2005), or the state's growing reliance on

statistical information, partly so that policy is seen as rational and accountable. Indicators can be used for descriptive, diagnostic, or predictive purposes, and consist of two major types: single indicators, such as the number of homeless people, or composite indicators, such as a housing affordability index.

City benchmarks, scorecards and competitiveness rankings rely on such indicators (ibid). Dashboards are thus essential to global city competition, which plays an important role in securing both financial as well as political capital (Rogerson, 1999; Robinson, 2006).

V. Opportunities and limitations of big data

Up to this point, this chapter has described the emergence of a new science of cities and provided a brief overview of big data and some of its sources, analytical tools, and applications. This section outlines the opportunities and challenges of big data, especially as they relate to the public sector. This will allow us to contextualize and further discuss the research findings presented in the following chapters.

Opportunities of big data

The transition from a data-poor environment to a data-rich one creates a number of new possibilities. It accentuates some of the benefits of the digital era: lower costs of storing and analyzing data, greater speed in doing so, improved transparency, and the promise of greater civic engagement. Big data can also provide more detailed, fine-grained information and so allow for more precise analyses. For example, it is now possible to collect significantly more data regarding urban mobility, at different scales and again, using faster and cheaper methods than before (Bibri & Krogstie, 2017).

Researchers can refine their understanding of formerly opaque phenomena, and can test assumptions quickly as monitoring & program evaluation can be targeted. It is also easier to conduct longitudinal studies and track trends instead of looking at snapshots.

As cities are understood as complex webs of relationships and information, big(ger) data can help us trace these networks more effectively as well. Where information flows present stumbling blocks to addressing longstanding problems, big data-enabled tools can provide powerful support.

Rogge et al (2014) have identified a number of specific benefits for public institutions as well. As part of a broader shift to digital, big data can accelerate the shift to a paperless environment and, perhaps more crucially, a data-driven culture. Government departments can share data better within and across jurisdictions, helping address the longstanding issue of collaboration in government (DeSouza & Jacobs, 2014). Greater monitoring capabilities can be used to improve the security of government systems and institutions as well. In particular, better internal monitoring can help identify issues earlier and reduce the risk of failure or corruption.

Citizen services and civic engagement are often heralded as major beneficiaries of the shift to e-government and big data. Improved data collection can help public agencies understand citizen needs in more detail and tailor services accordingly as well. Through the use of open data, the big data revolution can support civic engagement and foster innovation with "Government 2.0" (Chun et al, 2010), tapping into what Charles Leadbeater calls "the civic long tail" (2011), echoing Anderson (2004). Prediction markets and sentiment analysis can be leveraged to engage citizens and help them share their preferences as well.

> "In addition to its faulty theories, the urban science of the 1890s and the 1960s was hamstrung by three fundamental limitations: a lack of data, immature science, and limited computing power." (Townsend, 2015, p. 209)

By 2018, the big data revolution has ensured that limited computing power and lack of data are no longer the key challenges (or that they will not be for much longer). This perspective can be problematized, however, by recognizing that political and theoretical challenges have been as much an obstacle to progress as scientific and technological ones. In addition, while the big data revolution is here, like the rest of the future, it is not evenly distributed. In South Asia, as we will see in the following chapter, access to technology may have been addressed but there are still severe data limitations.

Limitations of big data

As a new field, big data has a number of limitations, beyond the existing ones regarding the use of data in urban planning and policy. Thakuriah et al (2017) identify four sets of disparate challenges: technology, methodology, theory & epistemology, and political economy.

Challenges	Characteristics
Technological	Urban information management challenges: 1. Information generation and capture 2. Management 3. Processing 4. Archiving, curation and storage 5. Dissemination and discovery
Methodological	1. Data Preparation Challenges (a) Information retrieval and extraction (b) Data linkage/information integration (c) Data cleaning, anonymization and quality assessment 2. Urban Analysis Challenges

	(a) Developing methods for data-rich urban modeling and data-driven modeling (b) Ascertaining uncertainty, biases and error propagation
Theoretical and epistemological	1. Understanding metrics, definitions, concepts and changing ideologies and methods to understanding "urban" 2. Determining validity of approaches and limits to knowledge 3. Deriving visions of future cities and the links to sustainability and social justice
Political economy	1.Data entrepreneurship, innovation networks and power structures 2. Value propositions and economic issues 3. Data access, governance frame-work and provenance 4. Data confidentiality, security and trust management 5. Responsible innovation and emergent ethics

Table 3. Extracted from Thakuriah et al (2017).

One of the foremost issues is that big data is not necessarily exhaustive and representative, especially with regards to user-generated data. The issue of the digital divide, or unequal access to the Internet and related technologies, continues to bias the data that is collected. In addition, big data often fails to meet the needs of analysts, especially regarding data quality (Mayer-Schoenberger & Cukier, 2013; Kitchin, 2014b). For example, while conventional travel surveys have many limitations, the data shared by ride-sharing giant Uber through its Uber Movement initiative is also deficient: for one, it does not track the all-important characteristic of trip purpose (French et al, 2015; Poon, 2017). Data quality is such a problem that data scientists spend the majority of their time simply cleaning and preparing data (Press, 2016).

Once the data is collected, it is frequently the case that the technology and algorithms needed to process it are not ready yet (harkening back to the challenges faced by PSSs over the last few decades). Most existing algorithms were designed for small data sets, and cannot be used for urban big data (Hashem et al, 2016; Bibri & Krogstie, 2017).

Furthermore, while big data holds the promise of deriving value at lower cost, there can be significant costs associated with the data management and analysis functions, both in terms of one-time capital costs such as purchasing technology and building the appropriate data infrastructure and ongoing operational costs including personnel, training and maintenance (Sivarajah et al 2017). The lack of trained professionals to handle big data, especially those that are also trained as urban planners, is another challenge in and of itself.

In addition, there are security concerns associated with storing and processing sensitive data using online tools, unrelated to personal privacy concerns. Corporate confidentiality can be risked, critical public information can be leaked, especially as distributed data on the cloud can be vulnerable to hacking and malware (ibid).

Beyond issues related to data itself, there are important limitations to what can be done with data, especially for planning purposes. Urban planning is, famously, a 'wicked problem' - a complex social problem that has no definitive or objective resolution, and cannot be 'solved'. As Rittel & Webber, the originators of the concept of wicked problems, stated: "The kinds of problems planners deal with - societal problems - are inherently different from the problems that scientists and perhaps some classes of engineers deal with" (1973, p. 160). Four decades later, Bettencourt (2014) invoked this to declare that, due to the inherent complexity[2] of cities, "planning the city is computationally impossible in all but the smallest towns" (p. 9).

For public administrators, big data is a challenge primarily due to the lack of policy direction and institutional capacity (DeSouza & Jacobs, 2014; Rogge et al, 2017). To begin with, a lot of data is localized in departments with differing jurisdiction, and there is little incentive to share data,

[2] The classic example provided by Glouberman & Zimmerman (2002) is that baking a cake is simple, sending a rocket to the moon is complicated, and raising a child is complex.

especially if departments are competing for funding. Even if incentive were developed, there is usually no clear mechanism to share data effectively. In addition, there is no guidance on how to use big data, at the institutional as well as individual level - most public servants are not trained to use big data. Policies regarding the ownership, confidentiality and security of information would also likely have to be substantially revisited in the age of big data.

Given that big data is not immune to existing challenges that data faces, it is wise to treat it with caution as well. Measurement errors, inconsistency, and skewed or missing data sets are liable to occur with big data as well. A much-touted example of using big data to address a public need was Google Flu Trends, which promised to predict flu prevalence earlier than the US Centers for Disease Control and Prevention. While it performed well at first, it eventually failed, in part due to spurious correlations with search terms (Lazer et al, 2014).

Similarly, promoting civic engagement through open data is fraught due to the digital divide. DeSouza & Jacobs (2014) also note, in response to the claims around citizen engagement, low engagement is not due to a lack of technological solutions, but due to other weaknesses in the democratic polity.

Dashboards are open to some of the same critiques, especially regarding the quality and accuracy of the data they present. They have also been criticized for promoting a narrow, overly instrumental, and often overly simplistic view of cities. They do not simply display facts and reflect the state of the city, but frame the city in a certain way. Dashboards are not just a neutral toolkit, but also complex 'data assemblages' with embedded politics in their own right (Kitchin, 2015). For policymakers, a decision support system is only of limited utility if it cannot help derive causality and can only offer correlations (DeSouza & Jacobs, 2014).

It is easy to get caught up in the hype surrounding big data, and forget about 'small data', whereas both are needed and are arguably complementary (boyd & Crawford, 2012). This also opens up the risk of paying attention to correlations and ignoring causality; ironically, the haste to apply technology can undermine scientific fundamentals (DeSouza & Jacob, 2014). It is especially problematic to focus on correlations when examining wicked social problems, as in the Google Flu Trends example described above. In the enthusiasm to apply data-driven methods, it is possible to use data for purposes it was not intended for, fitting square data into round holes.

The issues highlighted here point us to the need for a critical data studies (Dalton & Thatcher, 2014) where a number of underlying assumptions regarding big data are questioned. This is not only to puncture the bubble of big data hype but to ensure that there are reasonable expectations and that big data is pursued for the appropriate reasons. There is a need to interrogate the data assemblages, the structures and elements that constitute Big Data and its associated industries (Iliadis & Russo, 2016).

Contrary to the proclamations and hopes of many big data proponents, data is never objective or neutral (boyd & Crawford, 2012; Kitchin, 2014b). It is always produced through institutions with embedded values, using processes developed by people with their own biases. Bias is also inevitable when data is being interpreted; after all, "numbers have no way of speaking for themselves", as Nate Silver said (Marcus & Davis, 2013). The notion that data can speak for itself (and that it does not necessarily need to be interpreted in context) can be dangerously reductive and lead to simplistic interpretations (Kitchin, 2014b).

Finally, big data is accused of being a largely corporate phenomenon, driven by technology companies that profit off data (Dalton et al, 2016). It is certainly a moneymaking endeavour:

revenues from big data and business analytics exceeded $150 billion in 2017 (IDC, 2017). This profit motive can distort incentives around the collection and analysis of data, and is linked to the issue of representation and the production of 'data silences', where data that is not potentially profitable is not collected.

These limitations of big data are equally applicable in Asia, if not accentuated. Dynamic economic and population growth, including rapid urbanization, persistent widespread poverty, the postcolonial and post-socialist context, and variable political stability and state capacity mean that data is often questionable and fraught with tension. Stevens et al (2017) draw attention to the following six areas that require further attention in the Asian context:

1. *representation*: the digital divide is relatively deep in Asia, where much of the world's offline population resides (UNESCAP, 2016)
2. *unequal distribution* of the benefits and risks of big data: the poor are prone to marginalization, as big data often represents corporate interests
3. *history*: colonial and other historical data collection practices affect how local populations perceive modern data collection practices; in China, for example, big data is necessarily an instrument of the regime
4. *infrastructure*: the structural conditions under which data is collected, and how it is stored
5. *cost-effectiveness*: big data may not be necessary when simpler, cheaper proven approaches have not been implemented
6. *transparency and civic engagement*: far from empowering citizens and sharing more information, big data can be used to exert greater control where democratic institutions are weak

What we can expect of big data

To sum up, while big data offers unprecedented opportunities, there are important limitations to its potential. It is not perfect, and it is especially critical to remember that there is no such thing as raw data, and that all data is always manufactured. Big data is not *necessarily* transformative and will likely not wholly replace the existing uses of statistics and quantitative methods in cities, but can be complementary (Florescu et al, 2014).

Like any tool, big data does not solve any problem in and of itself, but it can help planners better understand and improve efforts to try and address problems (Bettencourt, 2014). The comparison between conventional travel surveys and Uber Movement data is instructive in this regard. While the Uber Movement data is not radically different from existing data that transportation planners already use, and suffers from biases as only Uber users are represented, it represents high trip volume and can help infer patterns (French et al, 2015; Poon, 2017).

Big data thus adds a powerful set of capabilities and tools to quantitative social sciences. This does not, however, preclude qualitative ways of knowing, or impose an empiricist epistemology. Critical GIS and radical statistics provide a framework to apply quantitative methods while acknowledging "situatedness, positionality and politics" (Kitchin 2014 epistemologies).

C. Perspectives on big data in Pakistan

Big data is yet nascent in Pakistan, with a limited track record so far. This section describes perspectives from interviewees and shares some recent developments. Specific challenges to the application of big data to urban planning in Pakistan are discussed in the following chapters.

Of my interviewees, all were familiar with the concept of big data to varying degrees. Three of them were expert data science practitioners who trained others in big data. Others had worked with technology for years and so understood big data quite well, even if they may not have been familiar with specific analytical tools such as Hadoop or Apache Spark. Interestingly, the two people who had the weakest understanding of big data were an urban planner and a senior civil servant working on provincial IT policy. Their understanding of big data was a vague notion of data generated through electronic devices and the Internet.

One of my interviewees, who works in the private sector providing data analysis services, spoke about two common misconceptions regarding big data in Pakistan: that it is simply about volume, and that 100 terabytes represents huge volume, and that it is generated through mobile phone applications. To him, the tipping point regarding volume (which, as defined above, is relative to processing ability) is that big data requires specialized big data architecture. He also highlighted that most mobile phone applications developed in Pakistan collect only a few data points: name, phone number, email address; hardly enough data to merit being called big data (even if it were in the order of millions of records).

A poor understanding of big data is clearly reflected in the "Big Data Analysis Reports" produced by Gallup Pakistan[3], which are simply presenting highlights from government reports (Gallup Pakistan, 2018).

The overarching view on big data was that there is very little of it. Outside of a few initiatives and organizations, big data does not exist as there is very little architecture that captures big data. Even when data is collected, it is kept in silos; whether it is a factory that captures structured

[3] Not a subsidiary of Gallup International

production data and does not integrate its video feed or other sensor data, or whether it is different government departments such as land, vehicle and tax records that are all separated.

This does not mean there is not enough data - but that it has not been collected and used in a certain way. Two overarching challenges identified here are the lack of data integration protocols or standards, and the lack of data literacy and trained human resources that can handle big data. This limits the ability of organizations to share as well as harness data. There was a common belief that even if someone were producing big data in Pakistan, they would never share it with anyone else. Unwillingness to share data was repeatedly highlighted as a challenge.

Most of the adoption of big data has been driven by banks and telecommunication service providers, which have also been among the fastest growing sectors of the Pakistani economy over the last decade (Ahmed & Ahsan, 2011). Applications include predicting if customers will default on loans from banks, or when customers will leave a mobile service provider (known as "churn"). A number of organizations have begun installing big data architecture, but are still struggling to derive value.

> "There's a lot of confusion in Pakistan, even with large corporate entities where you find 20,000 people employed and have their entire innovation departments. They have their own data science teams and they still don't know what they are doing." (Interview, May 12, 2018)

One of the first examples of big data in government in Pakistan is the development of the National Database and Registration Authority (NADRA) database that stores citizen identities. Teradata, a global data services provider, invested in Pakistan in 1999, from where it also operates its business in Afghanistan and Bangladesh (Alam, 2012; Khan, 2017). Teradata helped

NADRA develop an advanced bio-metric database that holds over 120 million identities and helped establish big data architecture that is used to this day (Malik, 2015).

A modern example is the Punjab Safe Cities initiative, which is installing cameras, microphones, sensors, and a dedicated fibre-optic cable across six cities in Punjab to create a monitoring and surveillance network. The system is operational in Lahore; a deeper examination is provided in the appendix. According to a manager, the Lahore Safe City collects 25-30 petabytes of data in one month alone.

All of my interviewees, bar none, identified a number of opportunities for big data and claimed that it would be helpful. Urban planners highlighted the availability of more granular and detailed data as a positive, as well as the ability to integrate and analyze disparate sets of data. One planner also reflected that data captured more recently and at higher frequencies might be used for trend analysis and reveal new insights regarding a phenomenon, especially as he has had to work with old data for much of his career.

Despite this enthusiasm, a common thread was that big data may not merit investment at present, when existing data is not even being used well. A few interviewees said it is too early to work on big data in Pakistan. Expensive technology investments may not yield results if people are not able to use it properly and could potentially reduce the pool of funds available for policy priorities such as education and health.

At senior levels of government, it seems that the consensus is that Pakistan must develop institutional capacity for big data. In May 2018, the Ministry of Planning, Development and Reform and the Higher Education Commission launched the National Center for Big Data and Cloud Computing, allocating Rs. 1.5 billion (approximately USD $12.25 million) to establish

twelve centers at universities across Pakistan (InnovatePK, 2018). The funding will be used to hire 50 faculty researchers, 100 graduate students, and 200 research associates. Five of the twelve proposed labs will specialize in using data to address classically urban challenges such as traffic and crime.

The federal Ministry of Information Technology and Telecommunication (MOITT) also passed a Digital Pakistan Policy in 2018, which has a section on cloud computing and big data. The policy calls for integrating federal and provincial government databases, removing legal and administrative barriers to sharing data for citizen service delivery, providing infrastructure, build skills, and provide measures for data security (MOITTa, 2018). In July, MOITT also released a draft Personal Data Protection Bill to protect personal data (MOITTb, 2018).

Chapter 3: The Use of Big Data in Urban Planning in Pakistan

This chapter shares research findings regarding the use of data in Pakistan. It begins with a profile of the Urban Unit, the provincial planning organization that is the paper's primary case study. This is followed by perspectives on the role and use of data in urban planning and policymaking in Pakistan, based on interviews conducted in May 2018. The chapter then ends with a section describing how data is used at the Urban Unit. While this chapter provides an overview of the different challenges identified by interviewees, the following chapter describes these challenges in greater detail.

A. Introducing the case study: Urban Unit and its role in Punjab

> 'technologies do not function independently of their environments, rather, they gain meaning only as individual staff members in a particular cultural and organizational context interact with them'. Campbell (1995, p. 104)

I. About the Urban Unit

The Urban Sector Planning and Management Services Unit Pvt. Ltd, usually shortened to the Urban Unit, is one of Pakistan's only dedicated urban planning firms, with a significant track record of conducting and coordinating urban planning and policymaking, concentrated in Punjab. It was established in 2006 as a Project Management Unit of the Planning & Development Department of the government of Punjab, and was converted into a private sector company in 2012, owned by the government of Punjab. It is governed by a board drawn from the public

service, academia, and private sector. Its head office is in Lahore, with satellite offices in a number of cities across the country.

The Urban Unit provides advisory, consulting, contracting and other services to the public and private sector in a wide range of areas relating to urban development, specializing in the use of technology. It has conducted projects across the country with over 35 national and international partners such as the World Bank and the Asian Development Bank. It employs over 400 people distributed in over a dozen sectors led by senior specialists, including urban planning, transport, water supply and sanitation, economics, environment, geographic information systems, and institutional development.

In practice, Urban Unit operates as a consulting and research firm that helps implement projects, either directly providing services to government departments, or acting as an intermediary between international consultants and the government. For example, it executed the Punjab Cities Governance Improvement Project, an initiative to strengthen the governance of Punjab's five most populous districts. It has also organized a number of events such as the Pakistan Urban Forum, a national week-long conference on urban issues, and created a platform to publish research in the form of the Pakistan Journal of Urban Affairs.

The Urban Unit is marked by its use of technology and promoting a data-driven approach to urban planning and analysis. It is known as a data hub in the Punjab government and has an exceptionally strong GIS development team, the best in the country according to several interviewees. What distinguishes their GIS and IT development teams is that their employees have domain and policy knowledge as well as having technical expertise, which allows them to interpret and respond public sector needs more effectively than other IT solution providers.

One way of understanding the Urban Unit's role is to think of it as a niche software developer for the public sector, specializing in spatial technology products. The table below provides a partial list of some of the projects that it has completed. Notably, nearly all of them feature a dashboard.

Project	Client	Partner
Annual Development Plan Monitoring Dashboard Sindh	Finance Department, Govt of Sindh	
Census of Brick Kilns in Punjab	Labour and Human Resource Department	School Education Department
Gender Management Information System & Punjab Gender Parity Report 2016	Punjab Commission on the Status of Women	DAI EDACE
GIS Based Computerization and Automation of Urban Immovable Property Tax Punjab	Excise Taxation & Narcotics Department, Govt of Punjab	
GIS Based Computerization and Automation of Urban Immovable Property Tax Sindh	Finance Department, Sindh	Excise Taxation & Narcotics Department, Govt of Sindh
Monitoring Dashboard for Road Sector in Punjab	Director General Monitoring & Evaluation	Communication & Works Department
Asset Management of Pakistan Railways Land by GIS based Computerization	Pakistan Railways	
Information of District Level Pension Management System		
GIS Based Land Record Management Information System		
Management Information System for MPDD		
GIS Based Data Analysis in Education & Health (SNG)		

Table 4. Extracted from Urban Unit, 2018.

II. The Urban Unit and the Punjab government

Understanding the Punjab government and its current political moment is key to understanding the organization. As the Urban Unit was established by, still receives funding and contracts from, and works with the government, it is necessarily beholden to government and political interests.

Punjab is, politically, arguably the most important province in the country. This is not just because of its demographic and geographic size, but also because it is widely seen as the stronghold of the Pakistan Muslim League-N (PML-N), the party that recently completed five years in power. Punjab was governed by the PML-N since 2008, and the provincial chief minister for the last decade was Shahbaz Sharif, the brother of ousted three-time Pakistan Prime Minister Nawaz Sharif. In the 2013 national elections, it was their dominance in Punjab that provided the PML-N with the ability to capture a majority of the National Assembly and thus broad electoral mandate. In the lead up to the 2018 elections, the PML-N faced strong opposition from rival parties such as the Pakistan Tehreek-e-Insaf (Pakistan Movement for Justice). As a result, the success or failure of the PML-N as a political force rested upon their success or failure as the government in Punjab.

Owing to these political incentives, chief minister Sharif implemented both headline-grabbing projects, such as the construction and launch of bus rapid transit and mass transit in urban Punjab, and ambitious governance reforms in health, education and planning. The speed and scale of these efforts have been praised by a number of international observers; a Chinese delegation remarked that "Punjab Speed" had surpassed the famous Shenzhen Speed, a label that the PML-N has happily adopted for marketing efforts (The Nation, 2017; O' Donnell, 2013).

Since its founding, the Urban Unit has been closely involved with the successful delivery of these initiatives. In particular, it has been a key stakeholder, along with the Punjab Information Technology Board, in the PML-N's efforts to introduce evidence-based policymaking and e-governance in Punjab.

One of the major efforts of the Punjab government has been to transform the economy and set the province on a path for sustainable long-term growth. The Urban Unit has been tasked with developing the Punjab Spatial Strategy, a 30-year spatial economic development strategy for the province. By identifying key growth nodes (such as economic clusters) and corridors, the Punjab government hopes to attract and coordinate public and private investments and thus align future economic and spatial policy. This entails gathering a significant amount of data from different departments and developing GIS instruments to provide a 'spatial portrait' of the province.

The Urban Unit is responsible for data collection, analysis, technical expertise, coordination, project reporting, and implementation. Given the magnitude and importance of this task, the Urban Unit assumes critical importance in the fortunes of the province's major political party.

III. The Urban Unit's positionality

Two visual artifacts of the Urban Unit provide interesting insight into its culture. The first is a large mural on a wall outside the office's boardroom, which displays a passage from the Holy Qur'an, which reads:

> Say, "Are those who know equal to those who do not know?"
>
> (Qur'an, 39:9, Sahih International translation)

This is rather revealing. It is not unheard of for offices in Pakistan to display some verses from the Quran, but they are usually standard pronouncements praising Allah and the Last Messenger. For the Urban Unit to display this demonstrates an intentional choice to make a statement. Given the Urban Unit's reputation for leveraging data, the statement is a clear assertion of superiority and, crucially, power. It is also, necessarily, a claim that a certain kind of knowledge is superior to another - the kind that the Urban Unit specializes in, which is based on quantitative methods and gathered through a top-down process.

It is reminiscent of Foucault's explanation of how knowledge is linked to power:

> "We should admit rather that power produced knowledge . . . that power and knowledge directly imply one another; that there is no power relation without the correlative constitution of a field of knowledge." (Foucault, 1979, p. 27; cited in Flyvbjerg, 2002).

The second artifact of interest is the Urban Unit logo, which is centered around four pieces of a jigsaw puzzle. This points to the organization's role in connecting different pieces, which could refer both to connecting and integrating data, as well as to coordinating between different institutions and departments in the government to gather and use this data more effectively.

Figure 5. Urban Unit logo

These two artifacts, taken together, seems to betray a contradiction in the organization. In the first, Urban Unit announces it superiority; in the second, it demonstrates a collaborative function. This contradiction may help understand some of the successes the Urban Unit has had, as well as some of its ongoing challenges.

B. Data and big data in urban planning in Pakistan

This section begins to answer the research questions presented in Chapter 1, drawing primarily on interview responses. While data is highly valued 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. This also means that while there is appreciation for, and interest in, the value of big data, it is somewhat early to talk of big data in the context of urban planning in Pakistan. There simply isn't enough big data being gathered, and 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. This point is discussed in the following chapter.

I. The use of data in urban planning

> "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 has been conducted for Fort Munro, a hill station being developed as a tourist resort (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, the Urban Unit is using data to develop the Punjab Spatial Strategy mentioned above, 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.

Partly in response to the recurring challenge of lack of data, the Urban Unit developed the Punjab Cities Growth Atlas 1995-2015, which uses primary and secondary data to develop a portrait of urban expansion and change in Punjab over a period of two decades, the first such analysis conducted in Pakistan. It identifies the distribution of, and connections between, all 194 cities in Pakistan, as well as presenting data on the demographic and spatial expansion, and land consumption, of the top 50 cities in Punjab. This data has been analyzed to present indicators such as contiguous urban development, area growth rate, area composition (the proportion of areas classified as high, medium or low density), and the change in area composition since 1995. This data is presented for all 36 districts in the province (Urban Unit, 2018).

Two things are unique about the Punjab Cities Growth Atlas. The first is that it draws upon a number of disparate secondary data sources, including but not limited to PSLM, LFS and PDS. While the data from these surveys is usually available in aggregated formats, the Urban Unit was able to obtain micro-data and process it to reveal information such as rural-urban disparities. This was a laborious process, with 40 people working for months to process and clean the data.

A second element is the methodology used to mark urban extents in Punjab. Instead of relying on administratively defined boundaries, the team obtained satellite imagery and used an algorithm to identify clusters of contiguous built-up areas (marked by contiguous buffers, defined as having over 50% or more built-up land within one square kilometer). This methodology was inspired by The Atlas of Urban Expansion developed by UN-Habitat, the Lincoln Institute of Land Policy and New York University, which also inspired the creation of the Punjab Cities Growth Atlas.

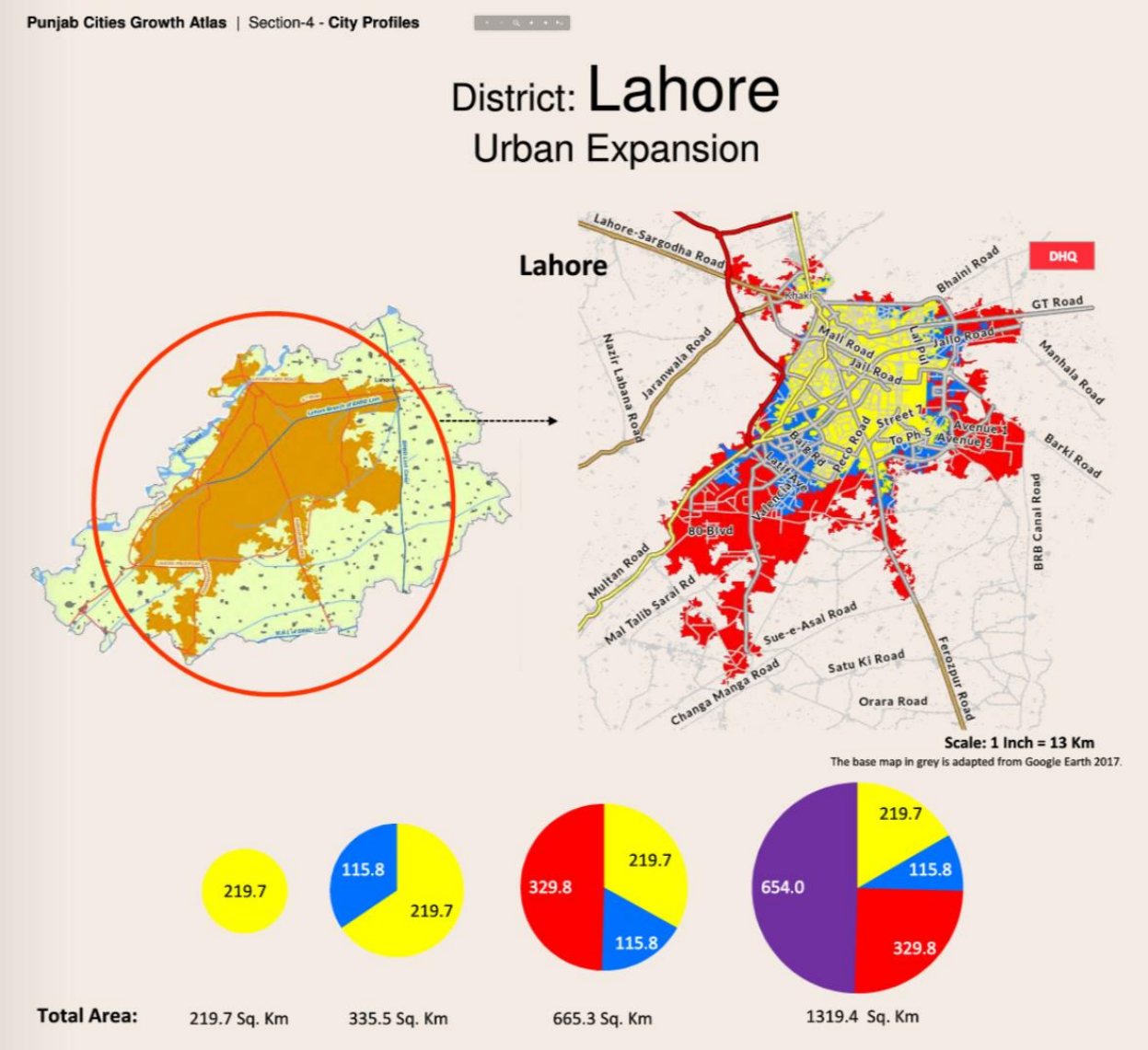


Figure 6. Section of Punjab Cities Growth Atlas – Lahore City Profile

The Atlas was only released in 2018 and represents a new knowledge product that can support evidence-based planning (Ahmed, 2018). At this point, less than six months after its launch, it is too early to say if it has been used by planners in Punjab or not.

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 at the Urban Unit 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, Urban Unit analysis 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, one cannot make generalizations. In the health sector, the government knows the number of patient facilities available in urban and rural areas, and the incidence of diseases. Recent outbreaks of disease may not be captured, and there may be a need to collect and process missing data using different methods. 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. For important projects, the government hires reputable and experienced people.

While I was waiting to meet this senior official, I spent several hours in an office for planning officers of a specific department. The office had seven desks, three on each side of the room with one in the center. Only four of the desks were occupied while I was there, with one person dozing and two others idly chatting to each other. None of the planning officers had computers in front of them, only small scattered stacks of files. I introduced myself to a planning officer and talked about my research project. He had a background in engineering and spoke about the application of data, but said that his section chief was out of the room as there was an interdepartmental meeting going on today. One of the points he made was that planning is mostly the domain of bureaucrats as politicians don't know anything about executing projects and development schemes. The second point he made was about how the bureaucracy was resistant to using data, and resistance to change in the bureaucracy.

Both of these public servants worked in the Planning and Development department, on the same policy file. The senior official claimed that data was being used reasonably well. The junior official did not have the resources to actually conduct any analysis and sat around idly all day long in the absence of his manager.

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. An interviewee highlighted this by talking about the budget:

> "Look at the budget – has anyone ever compiled budget data and talked about allocations and spending? A variance analysis every year doesn't teach you anything. If you say you went wrong this year, well, you've been off target for 10 years! *What are you doing about it?* That is the question." (Interview, May 12, 2018).

One of the things that became immediately clear in the process of interviews was 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.

Citizens are increasingly emerging as users of data. Most of my interviewees were familiar with the concept of open data and enthusiastic about opportunities for members of the public to access and use government data. The nonprofit Code for Pakistan was one of the first to introduce civic hackathons to Pakistan in 2013 (Usman, 2013), where software developers, designers and ordinary citizens come together and work on ideas for using technology to address social

challenges. The idea has taken off and hackathons are now organized regularly, such as Smart City: The Hackathon organized by Information Technology University in March 2018 (The Nation, 2018). The Urban Unit also organized a start-up challenge in August 2017; two of the four winners addressed urban issues (Urban Gazette, 2017).

These initiatives are limited, however, by the amount of data that is available. Several interviewees mentioned that there are a number of public servants who understand the value of engaging citizens and would love to encourage it, but are restricted from taking initiative and especially from sharing data. An alternative perspective is that if citizens are interested, they should take initiative and try to extract useful insights out of existing data that government has already shared and demonstrate that they have the appetite and capacity to use more data, which would (or at least could) then convince government to share more data.

Another area of citizen engagement is data governance. As technology becomes ubiquitous, more information is collected about citizens than ever before. The police carry out surveillance using cameras (described in the Safe City case study below), and private companies like Careem (a ride-sharing service) and Banks collect large amounts of sensitive information through their mobile phone applications. Citizens are not educated about, or protected from, the risks involved in this environment. Until now there has been no comprehensive privacy legislation that can protect citizens if their data is misused or stolen; indeed, the MOITT has only just released a draft data protection bill (Ali, 2018; MOITT, 2018b). This is an area that citizens should be involved in, both to advocate for a policy, provide insights and perspectives, and also to provide legitimacy to any policy that is developed. The issue of policy is discussed further in the following chapter.

II. Opportunities of using data

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. In the long term, it was felt that this would lead to better lives for Pakistanis.

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 cameras installed by Safe City Lahore to gauge flooding in different parts of the city during the monsoon.

As described earlier, transportation planning is often seen as low-hanging fruit to apply big data to. Unsurprisingly, 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.

Some of the most interesting responses highlighted the additional benefits of using data effectively. One interviewee spoke about the opportunity to 'defeat a number of myths', such as the myth that reporting accurate population numbers could reduce the resources allocated to lower-tier governments by upper-tier governments. This idea emerges from the fact that the National Finance Commission (NFC) award, which the federal government uses to distribute taxes among provinces, is predominantly calculated using population[4]. The NFC formula has incentivized provinces to misrepresent data by overstating their population and may also act as a disincentive towards population control. My interviewee seemed to suggest that data could help provincial governments recognize that accurate reporting would not significantly hurt them.

Another notable opportunity identified by interviewees was greater transparency and civic engagement. They shared that the open data movement could help people understand how the government functions, how resources are allocated, and how to engage more effectively. Sharing data could also help people understand their rights and make better democratic choices.

[4] For much of Pakistan's history, provincial population was the only basis for distribution, declining to 82% of the weightage in 2009 (The Gazette of Pakistan, 2010; Mustafa, 2011).

III. Major challenges to using data

This subsection provides an overview of the challenges to using data identified by interviewees. These will be discussed in further detail in the following chapter. The major challenges can be divided into three categories: data challenges, literacy challenges, and institutional challenges.

The primary data challenge is simply that there is too little data that is collected. There are not enough government mechanisms to capture and collect structured or unstructured data. If data is collected, it is frequently unreliable, collected through faulty processes or using untrained researchers. Even when data is collected, regardless of its quality, it is often not shared, or not shared in appropriate formats and with the appropriate level of detail. This speaks to the lack of data standards mentioned in Chapter 2.

The lack of data literacy presents another set of challenges. First, there is little awareness of the value of data among policymakers and in the public service, both at the senior level as well as the junior level. Secondly, most public servants and planners do not have the sufficient level of training and expertise in statistics and using tools to manage and parse data. Junior urban planners often struggle to even use Microsoft Excel for basic data analysis.

This links to institutional and political economy challenges to using data. The wider lack of data and data literacy have contributed to a policymaking culture that is not evidence-based in practice. The Pakistan Bureau of Statistics is considered by a number of interviewees to be too open to political interference, unable to operate autonomously. It has not effectively played the role of a central data broker or clearinghouse, at least not to the extent it should, according to interviewees. Vested interests often prevent data from being collected and used effectively; there

is a perception that data could challenge existing authority and be used to displace jobs. In addition, as identified earlier, there are a number of incentives to misrepresent data.

Finally, policymakers are largely unfamiliar with and do not trust big data and tools like decision support systems. As multiple interviewees said, building the system is far easier than building the capacity to use the system properly. Indeed, a number of interviewees stated that it may be premature to try and use big data in Pakistan, and it would be much better if existing data were appropriately used.

C. Data at The Urban Unit

This section describes the Urban Unit's capacity for data processing and evaluates the organization in the light of international best practices and the big data maturity model. There is a description of the Urban Unit's human resources, as well as how data is collected, processed, stored, and shared.

I. Human resources at the Urban Unit

With over 400 employees, the Urban Unit is one of the largest urban planning firms in the country. As an autonomous entity, it is free of traditional public sector constraints and can hire independently and set its own salaries. This has helped the Unit develop build up an impressive roster of qualified professionals, divided into a number of teams. For this paper, the urban planning, GIS, IT, IT infrastructure, and software engineering teams are the most relevant.

The urban planning team at the Urban Unit presently consists of four senior staff who handle urban planning and architecture, urban development, and research, and over a dozen junior research assistants. A major challenge for the team has been a series of exits of senior urban

planners, which has depleted the Urban Unit's institutional knowledge and planning capacity. This was evident to me during my internship at the Urban Unit in 2017, when two of the most senior members of the urban planning team left within a space of two weeks. It was also highlighted in interviews as a serious concern.

The firm's capacity in GIS and technology is what really sets it apart. There is one GIS team that handles spatial data gathering and cleaning, with separate GIS experts for land records, mapping, and transportation, along with a number of research analysts. The other team is the GIS development team, focused on building web and Android applications that use spatial data. The Urban Unit also boasts separate information technology and software development teams; the former focuses on process automations and is in charge of the Eyris project, while the latter develops custom MIS solutions. These teams collaborate frequently, taking on different roles based on project requirements. This combines a broad range of technical skillsets with domain expertise and creates an exceptionally strong base of talent that the Urban Unit draws upon for its different products.

These teams are backed by a large IT infrastructure team with over 25 employees. They handle 32 servers (with dedicated servers for GIS), seven enterprise storage systems and network security applications, and numerous computing devices that drive the Unit's work.

II. Data collection and processing at the Urban Unit

As data is core to its operations, the Urban Unit has developed a number of processes to handle data. It is worthwhile to mention that, despite the importance of data, the Unit has not formalized or codified these processes, relying instead on the vigilance of successive layers of management to ensure that data is being collected, analyzed and share appropriately.

Most of the data used is secondary, extracted from data provided by the Pakistan Bureau of Statistics and other agencies. The Unit also carries out primary data collection as needed, using both remote sensing methods as well as land surveys and mobile app-enabled data collection. This collection is largely decentralized; in my experience, most of the data being used by the urban planning team was collected by the team itself, or acquired upon request. There is no set process for collection. In one instance, the urban planning team required road length data for different districts and got it from the GIS department after making a request; in another, a senior analyst had to manually process raw data to get what was needed.

Data processing is also largely decentralized, usually carried out separately by different teams for their separate purposes. A degree of centralization is provided by the fact that the Unit has only one full-time statistician who parses data based on different needs, but after that all processing and analysis takes place by different sector specialists.

The lack of centralization is accompanied by the lack of policies regarding the use of data. There is no policy that compels employees to store data on a central server, even though all computers connected to the company's local area network can access a local server. The only safeguard against the theft or misuse of data is in the company code of conduct, which includes clauses on information security and confidentiality (The Urban Unit, 2018b).

III. Evaluating the use of data at the Urban Unit

This subsection evaluates the Urban Unit's data management practices based on best practices identified by IBM surveys.

In 2014, IBM carried out two separate surveys on the use of data and analytics; the first identified the value of speed in generating competitive advantage through analytics (Finch et al,

2014), and the second identified how different organizations used big data and analytics (Marshall et al, 2015). Based on the first, Nott (2015) identified the following categories to assess big data analytics and maturity for organizations:

1. *Business strategy* - data is integrated into the business model and exploited for innovation
2. *Information* - data and information are used as assets to improve processes and engagement
3. *Analytics* - data is leveraged for optimization and prediction
4. *Culture & operational execution* - the use of data-informed insights is embedded throughout the organization
5. *Architecture* - appropriate information architecture and standards have been deployed to drive maximum use of data
6. *Governance* - policies and procedures are implemented to ensure data and information are appropriately governed

The Urban Unit clearly relies heavily on data and analytics as part of its business model. Indeed, knowledge products based on information are at the core of the firm's value proposition, and the growth of the business over the last five years demonstrates that the Unit is delivering value. Revenues have increased from Rs. 569m in 2013 to Rs. 1.2b in 2017 (Urban Unit, 2018).

What is not clear is whether information and analytics are consistently used to improve processes, engagement and predict trends at the Urban Unit. Somewhat paradoxically, while the Unit creates decision support systems for the rest of government, I did not come across any evidence of the Unit optimizing its own operations using data. I only came across a few

initiatives, such as installing a biometric system for employees to check in and out of the office. I was not able to learn if these are part of a broader strategy.

In terms of execution, the Unit has successfully created a culture where data and analytical skills are highly valued and used regularly. The Unit has amassed significant human resources to use a variety of data, and supported them with an IT architecture that is sufficient to handle enterprise data needs. This capacity is limited to using existing data, not big data, but this is more than sufficient considering the lack of big data that has been collected in the country so far.

Data governance is an area of distinct weakness, with few policies in place. The Unit is exposed to risks due to a failure to integrate security practices which have been used in organizations around the world for well over a decade (von Solms & von Solms, 2004; Saint-Germain, 2005). The current international standard for information security consists of an entire set of practices that integrate assessment, monitoring and controls as part of an information security management system (ISO, 2018).

The Urban Unit can thus be identified as a 'Striver', an organization that is striving to develop innovation using big data and analytics, according to the classification developed by Marshall et al (2015). This places it in the middle tier, neither a Leader that applies a structured approach to using data, nor a Struggler that is risk-averse and has no formal innovation processes.

DeSouza & Jacob (2014) identify two distinct kinds of public organizations along the data continuum presented in Figure 1 - those that have historically dealt with 'small data', and those that have a track record of gathering and storing 'bigger data'. Organizations in the former category have not structured their decision-making and programs around data collection and use, and have yet to understand what data they need and how they will use it. For organizations in the

latter category, the challenge is to marshal resources to use data to its full potential. Many chief information officers recognize that they have not done much with existing data. The Urban Unit belongs to the second category.

More pertinently, the state of the art in the public sector is moving from data to 'bigger' data, but perhaps not quite Big Data yet. The Safe City Lahore initiative represents the cutting-edge of the use of technology and collection of big data in Pakistan, but has not yet progressed to data mining, analysis, and collaboration.

Chapter 4: Challenges to Using Data Effectively in Planning

This chapter 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.

A. Data access and reliability challenges

This section describes challenges relating to data: the lack of data, the inability to access existing data, and the unreliability of data that is available.

I. 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.

The Urban Unit also faces this problem. For example, two interviewees spoke about the challenge of getting 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). A senior manager tasked with the Punjab Spatial Strategy, expressed frustration with this:

> "It is very difficult, at the moment, for us to say something about the economy of a city, which is very important5. If you say that cities are the drivers of economic growth but don't know the size of the economy of the city and the ingredients of that economy, you cannot predict anything." (Interview, May 24, 2018)

Many of the data sets that the Urban Unit uses, such as the Multiple Indicator Cluster Survey (MICS) which was used for the Punjab Cities Growth Atlas and the Punjab Spatial Strategy, do not provide a rural/urban breakdown for indicators. 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, 2018).

In addition, some data is simply not collected, such as housing needs. For the Growth Atlas, an Urban Unit analyst used the number of housing units from the 1998 census and demographic

[5] Even if such data were available, of course, it would only account for activities in the formal economy. My interviewee believed that the informal economy is much larger than the formal economy, especially in the city; Arby et al (2010) estimated that 20% of transactions in Pakistan took place in the informal economy.

estimates from the PSLM Survey in 2015 to estimate housing shortages in different districts of Punjab.

> "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)

In contrast, sometimes, data is simply 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[6] 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)

A similar issue exists for big data; there is very little capacity to capture and harness big data.

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.

[6] CDA is the Capital Development Authority, the corporation in charge of Islamabad's planning and development.

II. Data access

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; the Punjab Spatial Strategy team consisted of professionals who knew how to work with data and were thus able to cobble together enough data to develop the strategy. The Unit purchased micro-data from the Pakistan Bureau of Statistics, digitized Survey of Pakistan maps, and carried out remote sensing which was verified by field visits.

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. Interestingly, the Unit also has a reputation for making it difficult to get their data. This could be because data is seen as a profit-generating asset by the Unit. The Growth Atlas, for example, is only available for reading on the Urban Unit webpage, with no ability to download a PDF version, copy text, or search information from the 135-page

document. The Punjab Safe City Authority has said they are open to sharing data, as long as recipients sign a confidentiality agreement and access data only at PSCA headquarters.

To its credit, the Pakistan Bureau of Statistics (PBS) has posted clear guidelines for sharing data on its website (PBS, 2018). Previously, aggregate data that had already been compiled was provided free, and there was a charge for special requests. Government agencies, universities, and partners such as the United Nations and World Bank can still get data for free, including micro-data (pending approval). Semi-government and private institutions, international NGOs, and individual researchers and students are charged for different data sets, depending on their needs and location (local requests are cheaper than foreign requests). This practice adds financial sustainability to PBS operations, but could hamper access to information, which is the democratic right of all citizens. It is not clear how this policy links to the Right to Information policies enacted at the federal and provincial level.

III. Data reliability

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

This quote summarizes another key issue: data reliability. 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 ("Methodology - Pakistan Bureau of Statistics", 2018). The same definitional debate came to the

fore in the 2017 census, when it was revealed that urban areas were counted very differently for Karachi and Lahore (Zaman, 2017).

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)

One planner expressed the belief that big data would be more reliable and less likely to be skewed. This was based on an understanding that big data is often captured by machines, in the

form of sensors and mobile applications. This view is only partially true, as big data cannot escape the biases of the humans who program the machines. The reliance on digital technologies also means that big data frequently over-represents those with access to the internet.

B. Data and technological literacy challenges

Data literacy was a theme that emerged in my exploratory research, during my internship at the Urban Unit, 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.

I. 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 my experience, junior urban planners and recent graduates had limited skillsets. For the preparation of the Urban Atlas, the Urban Unit hired planners to act as junior research assistants and tasked them with extracting and consolidating data from different surveys such as MICS and PSLM. These research assistants made a number of basic errors, were unfamiliar with Microsoft Excel beyond a few basic functions, and did not grasp simple concepts such as percentages, forcing 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). There is a relatively small critical mass of senior bureaucrats who are beginning to value data, but even they have still not developed the skills to use it effectively. The government has made efforts to address this deficit with training programs, as data literacy is crucial to achieving the government's drive to encourage evidence-based policymaking.

Big data and its potential are still largely unknown to the government, which is part of a broader lack of technological literacy. The Safe City initiative is creating big data that could be used, but their data is currently not being used, especially for planning purposes.

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. The interviewee from the private sector claimed, in fact, that he was unable to find enough skilled people for his own company despite offering competitive salaries.

It should be noted that the discussion above does not refer to employees of the Pakistan Bureau of Statistics (PBS) or any other government agencies dedicated to 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. Interestingly, the consultant who wrote a report on the service structure of the PBS (when it was being formed in 2011) called for the establishment of separate Data Processing/IT and Technical Wings, which would recruit from the open market and pay employees separately from government pay scales (Younis, n.d.). The PBS maintains a Training Wing that enables professional development for its staff through seminars and courses (Statistics Division, 2017).

According to the World Bank's Statistical Capacity Indicator, which is calculated as an average of the three scores of Availability, Collection and Practice, Pakistan has a score of 75.56. This is six points higher than the average for lower middle-income countries like Pakistan, and almost two points higher than upper middle-income countries (World Bank, 2018).

II. Data literacy at the Urban Unit

The Urban Unit is an anomaly as it can boast a relatively high level of data literacy. Nearly everyone works with data in some fashion, especially research analysts and the GIS teams; the only real gap is that there is only one full-time statistician, and no demographers. Over time, most employees develop a good feel for data and can identify which data is useful and which is redundant. As there is little big data in Pakistan thus far, the Unit has not developed capacity to handle it yet.

The experience of the 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. They also added offline functionality as field enumerators doing parcel surveys for land could not always have internet access. The GIS team developed offline caches that were loaded onto smartphones and then linked to GPS.

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.

The strength of the GIS teams and the cutting-edge nature of their work has also helped attract and retain talent. It has helped the Urban Unit deliver on innovative projects and made the organization an attractive place to work. The Unit, in turn, has recognized its GIS capacity as a critical asset and has rewarded employees by offering high salaries and regular increments. The GIS development team in particular requires a unique emerging skillset in spatial information technology. According to the team lead, they need people with training in both IT and GIS. Most IT specialists lack the ability to think in spatial terms, and GIS specialists lack the ability to do software development. The department has, however, been able to sustain and grow a team, all educated and trained in Pakistan.

III. 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 an 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.

Despite the growing recognition of the opportunities of data and big data, there are not enough good data science educational options, especially at the graduate level. The lack of programs is directly linked to the lack of qualified teachers and skilled practitioners who can train students. As big data has not taken off in Pakistan yet, local expertise is limited, and teachers have to be sent abroad before they can teach data science in Pakistan. While online learning is an option, it needs a significant amount of time and dedication, and needs to be supplemented with hands-on experience, and those experiential learning opportunities are rare.

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 (ITU, 2018).

IV. 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 a decision support system like Eyris, 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, p. 161).

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. Public servants are largely unfamiliar with the concept of iterating to refine technology solutions, even though the Urban Unit provides ongoing technical support and a team dedicated to making enhancements and upgrades based on feedback.

Due to their familiarity with coding and technical literacy, software engineers are finding job opportunities in the growing world of data analysis. They ultimately struggle because of a lack of statistical training, business expertise, and domain knowledge. This underlines the need for appropriate training programs to boost data and technological literacy.

C. 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.

I. 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. The bureaucracy has always been a powerful force in Pakistan (Jamali, 2015), and its power has always been rooted in the practice of documentation (Hull, 2012, pp. 7-12). Indeed, the word 'bureaucracy' itself was coined as an insult to the practice of writing down everything (Raadschelders, 1998, p. 142). One of Hull's major arguments in his book, appropriately titled 'Government of Paper', is that *parchis* (chits), files, and other papers are artifacts that tell a story about political economy and mediate social relations. If influence has historically been denoted by *parchis* and post-it notes, removing them also means limiting the influence and power of public servants.

Vested interests in the status quo also 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. I interviewed the CEO of private sector company that has developed a loan default prediction algorithm for banks; according to him, their algorithm is much faster and more accurate than the existing protocols used by banks. As a result, the use of this algorithm threatens the credibility and need for the 20-person credit department of the bank that has been operating for decades.

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. For instance, Pakistan is one of the few countries where polio still exists, and so there are massive door-to-door campaigns under way to ensure that every child receives polio vaccinations. Data from the population census conducted in 2017 could be used to identify households where there are no eligible children anymore, and enable better targeting. This also implies that fewer health workers are needed, however. Public servants recognize that greater efficiency could risk their jobs, so there is an incentive to hide or misrepresent data.

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. The CEO mentioned that there was initially a great deal of interest regarding big data in the private sector; driven by hype, banks, telecommunication service providers, and even hospitals established innovation departments and data science or analytics teams. When they were not able to derive value, they abandoned big data entirely and now only reluctantly embrace projects after proof-of-concept is demonstrated, "otherwise the war was already lost".

II. A 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 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: in 2007, five different organizations, reporting to four different ministries, produced and collected spatial data (Ali, 2007). At the time, none of them shared their data online, not for technical reasons but due to the lack of legal and institutional provisions for doing so.

Data type	Organization	Ministry
Topographic	Survey of Pakistan	Defence
Soil	Soil Survey of Pakistan	Food and Agriculture
Geological	Geological Survey of Pakistan	Petroleum and Natural Resources
Remote Sensing	Pakistan Space and Upper Atmosphere Research Commission	Scientific and Technological Research
Land use	Pakistan Agricultural Research Council	Food and Agriculture

Table 5. Federal agencies producing spatial data in 2007. Extracted from Ali, 2010.

This was partially addressed in 2011, when the General Statistics Reorganization Act was passed, establishing the Pakistan Bureau of Statistics (PBS) as an attached department of the Ministry of Statistics, by combining three federal agencies: the Federal Bureau of Statistics, the Agricultural Census Organization, and the Population Census Organization (General Statistics (Reorganization) Act, 2011). Regardless, agency boundaries and jurisdictional issues mean that data standardization and sharing are still limited.

III. E-governance - Shift to the data state, shift to the digital state

As with technological literacy, the use of data entails a broader shift towards using technology in government. This brings in e-governance, which has been defined as the use of information and communication technologies by government to share information with and deliver services to citizens (Palvia & Sharma, 2007; United Nations, 2018). The lack of e-government services, as well as their low adoption, means that government cannot collect enough data using electronic means.

Pakistan has been making efforts to introduce e-government since 2002, both in the shape of introducing online services for citizens as well as improving government operations. Progress has been slow and halting at best, due to a number of reasons, most significantly the lack of

appropriate ICT infrastructure, policy, and political will. While there is significant demand for e-government services, usage is limited due to poor awareness, inadequate help facilities, technical issues, and privacy concerns (Ahmad et al, 2013; Haider et al, 2015).

On the United Nations E-Government Development Index, which has been carried out since 2001, Pakistan has consistently received a low score and underperformed relative to countries in South Asia as well as countries with a similar level of gross national income. The index is a composite indicator, averaging the values of an Online Service Index, Telecommunication Infrastructure Index, and Human Capital Index (United Nations, 2018). The chart below shows how Pakistan's scores compare with those of neighboring countries.

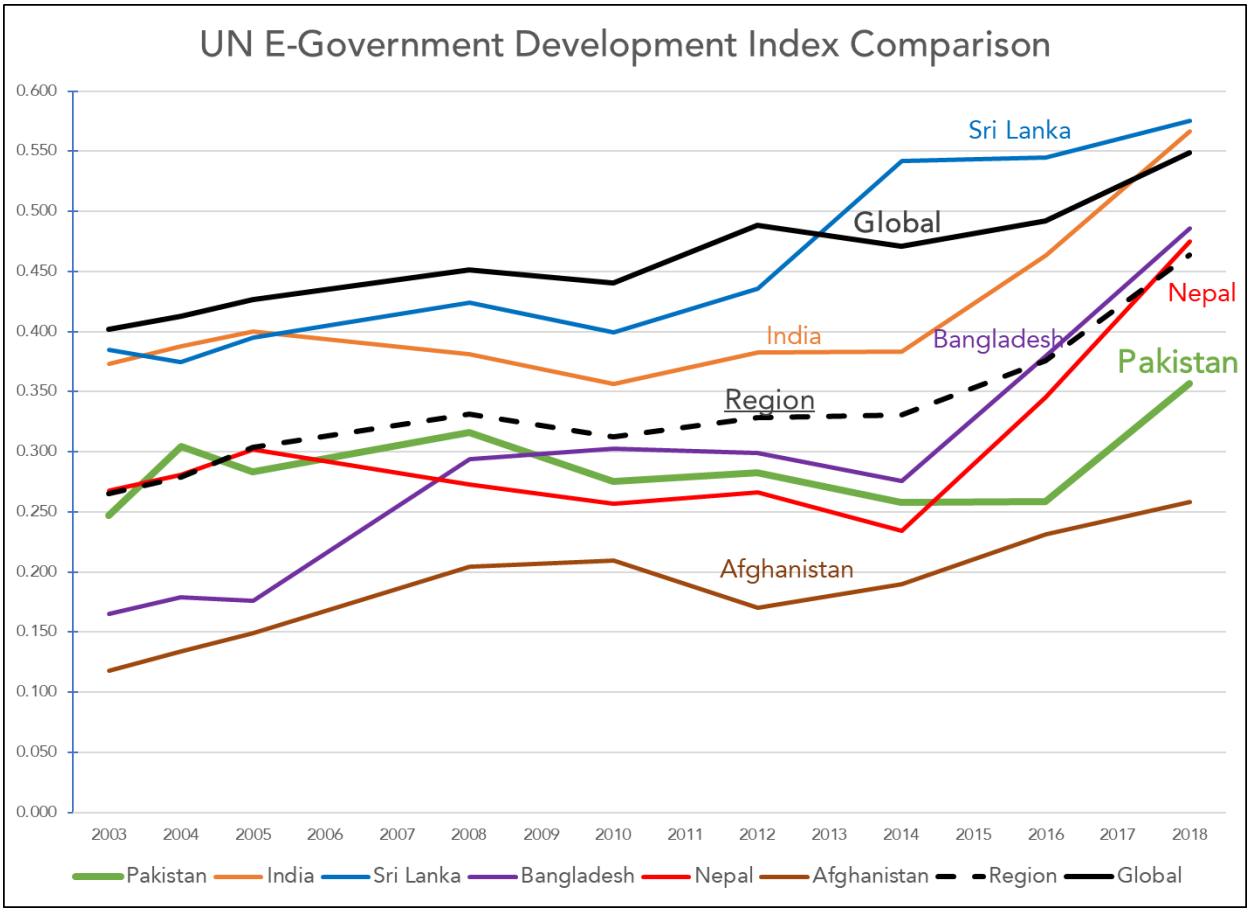


Figure 7. Pakistan's e-government scores. (Compiled from UN reports).

Pakistan's poor performance is reflected at the local level; the UN also carried out a pilot study of local online services in 40 cities around the world. The only Pakistani city selected, Karachi, was ranked at the very bottom, able to meet only 16 out of 60 indicators, including zero service provision indicators and only one participation indicator (ibid, p. 159).

Progress is being made. A closer examination of the scores reveals that Pakistan's low score is not due to a lack of online services offered, where Pakistan actually performs well. Rather, the issue is that people cannot access those services - Pakistan has poor scores for infrastructure (especially internet access) and human capital (reflecting education). The introduction of 3G and 4G in Pakistan in 2014 have increased internet penetration, and the number of broadband subscribers has gone up from 5 million in 2013-4 to 44.5 million in 2016-7 (Pakistan Telecommunication Authority, 2017). The government also claims that Pakistan will become the first country in South Asia to test 5G technology (Geo News, 2017).

One of the most successful examples of e-governance (and technology implementation) in the country is the National Database and Registration Authority, which is considered a model around the world for the successful implementation of a biometric national identity registration system. It boasts over 120 million citizens in its database, the largest and most detailed repository of citizen data in Pakistan (Malik, 2015). Beyond issuing national identity cards and other records to assist in identity verification, NADRA also shares its data with other agencies for functions such as compiling electoral rolls and lists of social service recipients (Bhatti et al, 2015, pp. 51-52). NADRA also implemented the first Safe City in Pakistan in Islamabad (NADRA, 2018a), and has implemented a number of projects in other countries as well, such as Nigeria, Kenya and Bangladesh (NADRA, 2018b).

Punjab is at the forefront of e-governance in Pakistan, with dozens of projects implemented by government agencies, especially the Punjab IT Board (PITB) and the Urban Unit, guided by the Chief Minister's Secretariat and the now-defunct Institutional Reform Group (IRG). In 2012, an IRG report identified 57 different "service delivery innovations" that had been implemented, in the areas of health, education, solid waste management, procurement, land records management, and law enforcement, among others. The Digital Punjab report published by the PITB claims that over 240 IT projects have been completed (Punjab IT Board, 2017).

In practice, not all of these innovations can claim either full implementation or success. An example is the e-Filing and Office Automation System (e-FOAS), an initiative to digitize and store government documents and schedule meetings more effectively. According to the Digital Punjab report, e-FOAS is being implemented in 13 different agencies and has increased efficiency (ibid, p. 64). In practice, when I visited the Planning & Development Department, I found that most desks had stacks of files and most employees did not even have computers. Files were still being processed by junior staff carrying them from senior officer to senior officer, who would glance at the files and then either sign or make a small comment. Ironically, I found a brochure for e-FOAS lying next to two tall stacks of files. My only meeting with someone in the department, who was actually someone working on IT, was twice postponed because the person's secretary had double-booked him. This was still very much the paper bureaucracy described by Hull (2012), and initiatives like e-FOAS are threatened by the very dynamics he described.

Photo of a desk at the Planning and Development Department, showing e-FOAS brochure.

The most prominent e-governance innovation in Punjab, arguably indeed in Pakistan, is the Citizen Feedback Model, a World Bank innovation that was first piloted by Zubair Bhatti, who was a District Coordination Officer in Jhang, Punjab at the time. The model is simple and powerful: citizens who visit a government office receive a phone call or a text message asking them about their experience (Vazquez-Dodero & Bhatti, 2013). Over 4 million citizens had received the service by 2014, and the model was a hit with citizens, public servants and politicians alike, being extended to different departments and shared internationally (Bhatti et al, 2015, pg. 51; Kunicova & Bhatti, 2015).

The model is especially relevant to this paper for its role in collecting data that is then used for monitoring and performance improvements, and loaded onto dashboards:

> "...data are collected and categorized daily, serving as a permanent survey and providing both a snapshot and a continuous stream of information about services throughout a province, a district, a subdistrict, or other local jurisdiction. The data help identify patterns and trends in the performance of departments and services...in effect a permanent survey seeking information from users of services. It prepares a live, electronic report card for every facility, as responses are associated with the facility that reported the cell number." (Bhatti et al, 2015, pp. 62-63)

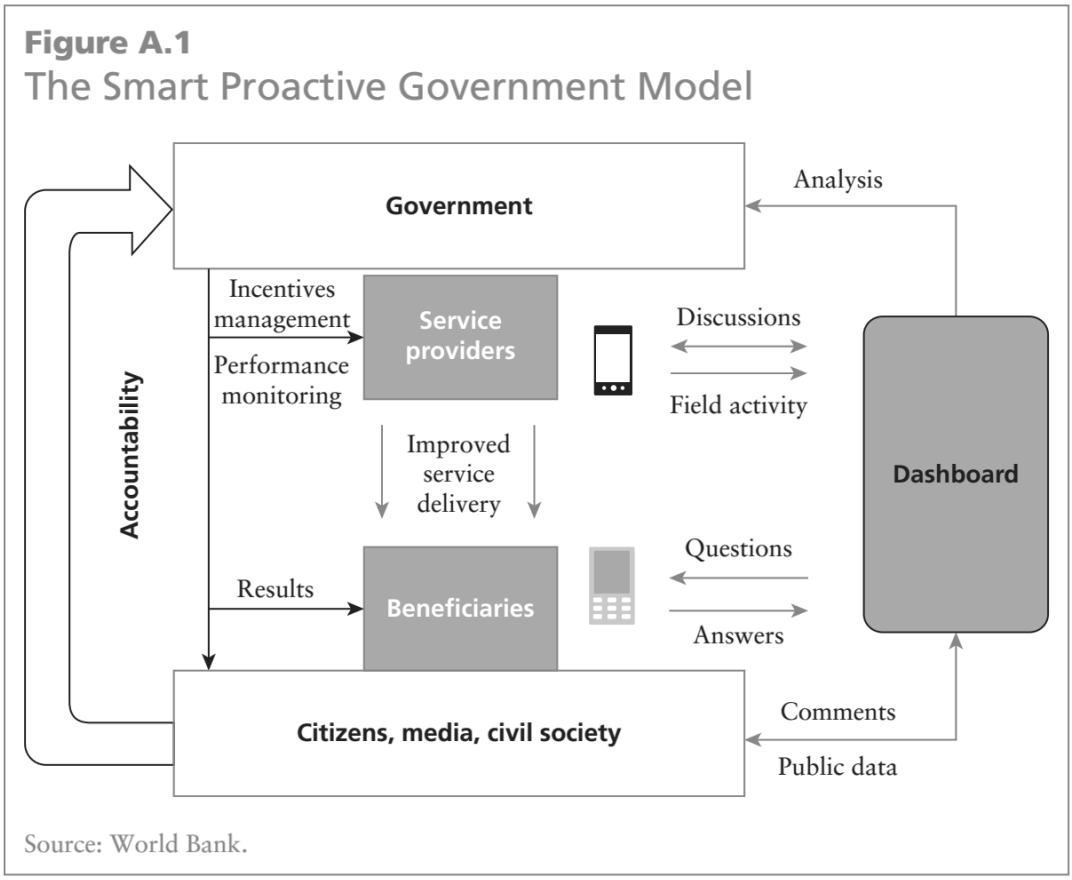


Figure 8: Diagram of the Smart Proactive Governance Model (Bhatti et al, 2015, p. 86).

The World Bank has since expanded this model into the "Smart Proactive Government Model", based on the use of SMS, call centers, smartphones, and dashboards (ibid, pp. 85-86).

IV. Policy gaps

An underlying issue with reforms to introduce evidence-based policymaking and e-governance has long been a lack of adequate policy. Previous efforts to use data have been patchy, relying on projects or enthusiastic bureaucrats, but were not backed up by the heft of official government policy. The lack of privacy policy or any data protection regulation is particularly of concern, especially as there have been reports that sensitive data from NADRA has already been compromised (Attaa, 2018).

It is only recently that the federal and provincial governments have approved policy frameworks regarding data and information technology. The federal Digital Pakistan Policy was approved in May 2018, Pakistan's first IT policy since the very first IT policy and action plan in 2000. In July 2018, the MOITT released a draft Personal Data Protection Bill for feedback. At the provincial level, the Punjab IT policy was passed in May 2018, one day before the government's tenure ended. These are the first government policy documents in Pakistan that have specific components to promote the use of data and refer to big data.

The federal Digital Pakistan Policy includes the following components: developing a framework for cloud services and data regulation; developing additional infrastructure, including data center clusters; improving data education and training, as well as placing dedicated IT human resources in government departments and focusing on capacity building for big data; implementing further e-government interventions and e-procurement; integrating federal, provincial, and

interdepartmental databases to enable data mining; and removing legal and administrative barriers to data sharing (MOITT, 2018).

The Punjab IT Policy emphasizes similar issues, such as bridging the digital divide by increasing internet access and data literacy programs, enhancing e-governance by expanding online citizen services and access to government information, and building a knowledge-based economy by launching an open data initiative, setting up data centers, and defining sharing guidelines.
The draft Personal Data Protection Act provides guidelines for processing, storing, sharing, and securing personal data, as well as defining the rights of data subjects, especially the right to withdraw data and ask for erasure of records. The Act also calls for the establishment of a National Commission for Personal Data Protection (MOITT, 2018).

D. 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. Pakistan has witnessed 'runaway urbanization' (Kugelman, 2014) across the country, mostly unchecked by regulation, a phenomenon that has not gone unnoticed by the public (DAWN, 2005; Raja, 2016; Randhawa, 2017). This is despite significant government efforts; urban policy has been an important component of national planning since Pakistan's first Five Year Plan in 1955 (Qadeer, 1996).

These national plans have largely focused on a limited set of issues revolving around housing and municipal service delivery, with little attention to spatial land use planning and development.

For the first 50 years of Pakistan's history, the focus of urban policymaking was narrowly focused and served to enhance private interests instead of public welfare, as Qadeer notes:

> The idiom of housing schemes and plot development forged in the preindependence tradition of town planning was consolidated in a policy instrument. And this has remained at the heart of urban strategy in Pakistan....Policy issues such as planning legislation and urban administration, traffic management remained unattended.
> (Qadeer, 1996, p. 452)

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[7]; 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

[7] Local governance can enable better monitoring and evidence-based planning in Pakistan (Mehboob, 2016).

the behest of international donors (Hasan, 2012b). None of these plans have been implemented in full, and many were consigned to filing cabinets almost as soon as they were developed.

The uncoordinated and fragmented nature of planning is a consequence of the 'perplexities' identified above, most importantly the absence of an overarching policy framework. There are no national or provincial spatial strategies that can serve as the connective tissue for city-specific or even district-level spatial plans. The Punjab Spatial Strategy being developed by the Urban Unit is the first of its kind in Punjab.

One of my most important and unexpected research findings was learning about exasperation with professional planning in Pakistan. 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.

Urban planning is established as a regulated profession in Pakistan, usually referred to as 'town planning' due to the legacy of British colonial planning. It consists of a closed circle of practitioners, regulated by the Pakistan Council of Architects and Town Planners (PCATP), which was formed by the PCATP Ordinance 1983. Only town planners and architects who are members of PCATP can perform planning functions. The field is dominated by architects: of the approximately 7,200 registered members of PCATP, less than a fifth are planners. PCATP effectively controls planning in Pakistan by tightly regulating membership through high barriers to entry and aggressively ensuring that non-members are not able to practice, as is clear from the disputes documented in their meeting minutes (Pakistan Council of Architects and Town Planners, 2018).

This gatekeeping has a number of implications. The membership and accreditation requirements limit the number of town planners and even certified planning university programs, especially considering the pace and extent of urban growth. For a country of 207 million, Pakistan has under 1300 registered town planners; for crude context, Canada with a population of 35 million has over 6,900 registered members of the Canadian Institute of Planners (Canadian Institute of Planners, 2017).

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" (Shah et al, 2007). The reliance on master plans, a mostly outdated approach, is itself emblematic of an outdated planning system.

Planning in Pakistan is also marked by a lack of public participation and involvement; for most plans and policies, consultation is nonexistent (Hasan, 2012b; Haque, 2018). This also limits the potential of any civic engagement opportunities generated through open data. Ironically, the lack of data should mean there is a stronger incentive to tap into local knowledge.

Finally, career growth opportunities for town planners are highly limited. Planners can only get jobs in development authorities or district councils, where there are limited positions, and can only move up two pay scales, from Grade 17 to Grade 19[8], before retiring (Interview, May 24).

[8] Civil servants in Pakistan are paid on a scale that consists of 22 grades (Finance Department, 2017).

Many planners are restricted to working for housing development societies, doing not much more than passing development approvals.

This has led to governments facing difficulties in recruiting experienced and skilled planners. During my internship, I observed an increasing reliance on outsourcing planning work to consultants. Outsourcing to independent and international consultants is part of a broader phenomenon in government, due to low in-house and local expertise, respectively. Its effectiveness is limited and questionable at best, increases costs, and limits accountability as consultants cannot be held responsible.

Chapter 5: Conclusions

This chapter concludes this Major Paper. It begins by sharing recommendations from interviewees and proceeds to outline future research directions regarding the use of big data in planning in Pakistan.

A. Recommendations for using data more effectively

> "There is a phrase in Urdu - *karon ga to badshah ki beti se, warna kunwara maron ga* [either I will marry the princess, or I will die a bachelor] – whatever data you have, use it, make something work, give it to an authority and maybe they will give you more." (Interview, May 12, 2018)

My final question in each of my interviews solicited recommendations on the effective use of data. This section summarizes the responses I received, organized into the same categories used for challenges: data access and availability, literacy, and institutions.

<u>Data use, access and availability recommendations</u>

1. **Share the data**: Several interviewees suggested the need to encourage a culture of data sharing, both within government and with the public, to encourage open data. They also mentioned the importance of sharing findings and insights gleaned from data, as these can spur productive debates and further knowledge creation. This recommendation aligns with the broader imperative for government departments to collaborate more effectively.

2. **Develop a common framework**: Another priority is to ensure that there is a common framework for data, with standardized definitions and a clear taxonomy to enable sharing, comparability, and interoperability.
3. **Work with what you have**: As the quote above emphasizes, it is important to work with what is already available, even if it is limited and unreliable. This can build the case for more data collection and sharing as the benefits become apparent.
4. **Start with low-hanging fruit**: Planners and researchers should recognize that data is political and that some datasets are more sensitive than others. To limit backlash, they should try to identify non-controversial issues to work on and demonstrate value there.
5. **Build a central database**: There is an urgent need to build a centralized, user-friendly interface and database to collect and share data. Pakistan needs a local equivalent of open data hubs such as Data.gov (United States of America) or Open.Canada.ca (Canada). Interviewees emphasized the user-friendliness of this system as of critical importance.
6. **Do a self-assessment and inventory**: A first step to effective data management and processing is to conduct an assessment of what data is already collected and what is needed. An inventory of data assets and needs would help set priorities for any new investments. (*Author recommendation*)

Data literacy

1. **Build data skills**: Planners need to be trained in statistics and must develop a foundational base of analytical skills that they can apply or expand upon as needed.
2. **Ask the right questions**: Planners need to develop a strong conceptual understanding of data - what it represents and what it can be used for. This will help them critically approach data and ask the right questions when they receive a dataset.

3. **Improve technological literacy**: As with data literacy, all planners and public servants must have a minimum level of competence with computing. They must be encouraged to learn how to use applications beyond Microsoft Office, which will develop familiarity and comfort with technology.
4. **Iterate to learn**: It is important to start using technology and not wait for it to be perfected; usage will help improve the quality of the tool. This parallels the earlier recommendation to work with what is available.
5. **Assess needs appropriately**: One interviewee bemoaned the waste of public funds in the name of promoting technology: "Buying an i7[9] just to make Word files is of no use. That is what is happening in our departments. They can buy a cheaper computer that will have the same life." (Interview, June 1).

Institutional challenges

1. **Empower the Pakistan Bureau of Statistics**: Interviewees appreciated the restructuring of statistical bodies in 2011 and encouraged further reforms to the PBS. They called for modernizing the agency and concentrating its power as a central body that works with and supports provincial bodies.
2. **New institution to manage big data**: Given the data revolution, some interviewees suggested that there should be a new institution created specifically to manage big data, starting with the management of digitized data. This institution would have a committee of experts drawn from the public service, private sector, and broader citizenry to consider and adapt international best practices to the local context.

[9] Referring to Core i7 computer processors, the latest and most expensive processors developed by Intel.

3. **Remove political interference**: One interviewee highlighted the need to ensure that the PBS and other statistical bodies could operate free of political interference and manipulation, especially considering the concerns raised regarding the population census.
4. **End the *daak***: *Daak* is the Urdu word for post and refers to the way in which orders and communication circulate through government offices in Pakistan. There is a longstanding reliance on physical files, which prevents transitioning to a paperless environment. As Hull (2012) described, though, the 'political economy of paper' is resistant to change and blocks digitization, transparency, and ultimately, evidence-based policymaking.
5. **Institute a legal framework for data**: At present, there are limited laws that govern public and private data. The government must develop an appropriate legal framework for data that covers collection, security, privacy, and authentication.
6. **Create an urban observatory**: An innovative suggestion, specifically relevant to encouraging the use of data in local government, was to create an observatory or lab at the city level to manage municipal data.

There is no simple solution. Sometimes new data is needed, sometimes new tools, sometimes new policy, and sometimes new people - likely a combination of each.

These recommendations echo previous ones that have been made in public fora and in the literature. For example, at a seminar on data gaps and evidence-based policymaking organized by Sustainable Development Policy Institute (SDPI), the key policy recommendations included strengthening links between statistical bodies, improving interdepartmental coordination, creating a national system for data sharing, promoting data literacy, and aligning definitions of indicators (Sustainable Development Policy Institute, 2015).

Similarly, a paper on harnessing big data calls for mutual respect between planners and data scientists, increased data availability, better tools and safeguards, and training for new and interdisciplinary skillsets (French et al, 2015).

Klosterman (2008) has suggested four principles for urban modeling that can extend to the broader use of data, especially through technology:

1) All models are wrong - some models are useful: all models are simplifications and are necessarily wrong, but they can still serve a purpose
2) Prediction is hard, especially about the future: planners and decision-makers cannot rely on forecasts or expect them to be accurate, but forecasts can be valuable for scenario planning, especially when assumptions are explicitly identified
3) Keep it simple, stupid: models are more likely to be used when they are simple and easily intelligible to users.
4) Use it because it's BAD: available information always suffers from imperfections, but it is also the Best Available Data. Take this into account when designing tools.

B. Reflections

This paper reveals that Pakistan's urban planners and policymakers are a long way off from using data effectively, let alone big data. There are a lot of pieces of the puzzle that still need to be put together, most importantly in terms of human resource capacity, enabling infrastructure and appropriate policy and legislation. The thorniest issues, as always, relate to power.

I. The politics of data

Data is political. This is well understood in Pakistan. It can help maintain an illusion or it can bring down an empire. There is, therefore, an urgent need to consider the political implications of

big data. Who uses it, and for what ends? Take the example of the loan default prediction algorithm described in the previous chapter. Such algorithms can inadvertently discriminate against people from lower income communities (Matsakis, 2018). This discrimination may not be intentional, but it is certainly real and malicious, as Virginia Eubanks (2018) has described.

Acknowledging the political implications of data allows us to engage with political barriers to the use of data. We must grapple with the possibility that sometimes data is not 'missing' by accident, but rather it was not collected, or that it was destroyed on purpose. It is not unusual to hear of fires breaking out in government offices on a regular basis in Pakistan; by some curious coincidence, these fires often seem to wipe out land records (Express Tribune, 2014).

The monitoring and surveillance infrastructure set up by the Punjab Safe City Authority (PSCA) also raises a number of worrying questions. Cameras capture personally identifiable data, and there are no guidelines or regulations regarding its use. The system tends to target the weakest, those who are homeless and cannot help but be in front of the camera. PSCA monitors consider loitering a crime, for instance. The Safe City panopticon represents a new technology of control.

New tools are often marketed as public service reforms but are also political projects. At the top of every screen being used to manage emergency dispatch services at the PSCA, there is a line that reads *'The Vision of Chief Minister Punjab'*. A large picture of Shahbaz Sharif can be found in the main video monitoring room. Clearly, this shiny innovation is being used to shine up political resumes as well.

As described in Chapter 2, dashboards represent complex assemblages of power and technology. In the case of the Urban Unit, these dashboards provide striking visual evidence of their capabilities, regardless of how well the dashboards are actually used and the impact of these

investments. Given what we now know about data and technological literacy in Pakistan, it is quite possible that the PSS implementation gap identified by Geertman (2017) could remain in Pakistan for years to come.

The drive to smart cities is also an act of branding and global city competition. Within a few years of the launch of Safe City Lahore, it was included in a brochure to demonstrate the safety of the city (LUMS, 2017).

II. No algorithm for better decisions

Both interviewees as well as the literature review in Chapter 4 suggest that the core issues with poor planning are not confined to data or a lack of knowledge, but institutional and political issues. While data can certainly help move towards evidence-based decision making, it can also be distorted for what some call 'decision-based evidence making'.

This phenomenon was immortalized in the TV show The Wire, where policemen who were ordered to bring the homicide rate down simply started classifying crimes in different ways, a phenomenon known as 'juking the stats' (Sheehan & Sweeney, 2018). This was the fictional depiction of something that took place in reality. In the early 2000s, the city of Baltimore received attention and awards for its data-driven performance measurement strategy, CitiStat. The system relied on the aggressive use of data and targets to improve service delivery, such as filling potholes. While CitiStat was certainly successful in parts, it also created perverse incentives (Behn, 2007). Journalist and author David Simon, the man behind The Wire, documented how the numbers were artificially twisted by inflating the crime rate of the past, making it seem that crime in the present had gone down dramatically (Simon, 2015).

This was a manifestation of Campbell's Law:

> The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor. (Campbell, 1979, p. 49).

'Evidence making' may also be taking place in Punjab, but at some cost. Shahbaz Sharif's government was famous for aggressively monitoring performance, but there were concerns that this may lead to falsified statistics (The Economist, 2018), as illustrated here:

> "Sharif has a 'heads will roll' approach. The idea is that if you put people under enough pressure, they will deliver," said a senior official and educationist working on the Punjab reforms, who asked not to be identified. "A district education official under that much pressure is bound to invent some numbers." (Naviwala, 2016)

III. Data can transform planning practice in Pakistan

There are two perspectives on the opportunities of using big data effectively in Pakistan. One set of opportunities is about the management of cities and about optimizing the functions of planning: the ability to track what is happening, derive new insights, create fine-grained plans, and operate far more efficiently than ever before. A different set of opportunities is about fundamentally remaking planning as we know it in Pakistan today.

Implementing the new science of cities can help policymakers and planners see cities in Pakistan not merely as agglomerations of labor and capital but as networks of social relations. By

observing interactions, planners can now pay attention to and analyse not just the built but also the social infrastructure of urban Pakistan and examine lived experiences in entirely new ways. Open data can be used to engage citizens in new ways.

While urban planners are being creative about using data, this is not enough. Short-term innovations can only act as duct tape in a fundamentally ineffectual planning system. Data can help bring about a transformation in how planning is practiced in Pakistan, not just by integrating data but by changing the very meaning and scope of planning.

C. Future research directions

This Major Paper barely scratches the surface of the research that needs to be done in this area. Different research directions can be identified for big data and smart cities, both in the context of the global South and without it. In this paper itself, there were a number of perspectives that could have been further explored, such as the role of data in enhancing urban sustainability, or the political implications of data collection in Pakistan. Another theme is the broader issue of public sector innovation and reform in Pakistan, or perspectives on the emergence of smart cities in Pakistan. A future study may examine different approaches to enhancing data literacy for public servants. Given that data sharing was such an important theme of this paper, more research is needed on public sector collaboration with regards to data and big data.

There were also a number of technological innovations that were not discussed, such as recent initiatives to digitize land records and urban properties. The Safe City initiative also requires much more attention, particularly using the lens of governmentality and control.

The use of data in urban planning implicates a number of ongoing trends in Pakistan, each of which merit further exploration and critical inquiry: urbanization, urban planning, local

governance and devolution, evidence-based policymaking, e-governance, open data, civic engagement, and citizenship. Each of these trends are mirrored by developments in other countries and other parts of the world. For example, there is room for an entire body of research examining the links between open data and urban entrepreneurship opportunities.

New research could emerge along the lines of different technologies (sensors, dashboards, smart grids, driverless cars), different industries (housing, transportation, water and sanitation, and energy), and different applications (modeling, machine learning, and more).

More generally, the use of technology in cities in the global South is a major research area that has not received sufficient academic attention. The vast majority of the research on the use of technology for urban planning and policy relies on examples from developed countries in the West. There needs to be more research on the range of technologies being used in the cities of the global South, how they are being used, and what forces are driving their adoption. More empirical work also needs to be done to assess the outcomes of this transition. There is also a need to closely examine the emergence and nature of smart cities in the global South.

There is a significant amount of theoretical work that can be undertaken as well. For example, it is important to more deeply explore how data is used to exert influence, and how big data could reshape power relations in different contexts, whether in a public sector organization, in a political coalition, or in the global South (building from Rogge et al, 2017).

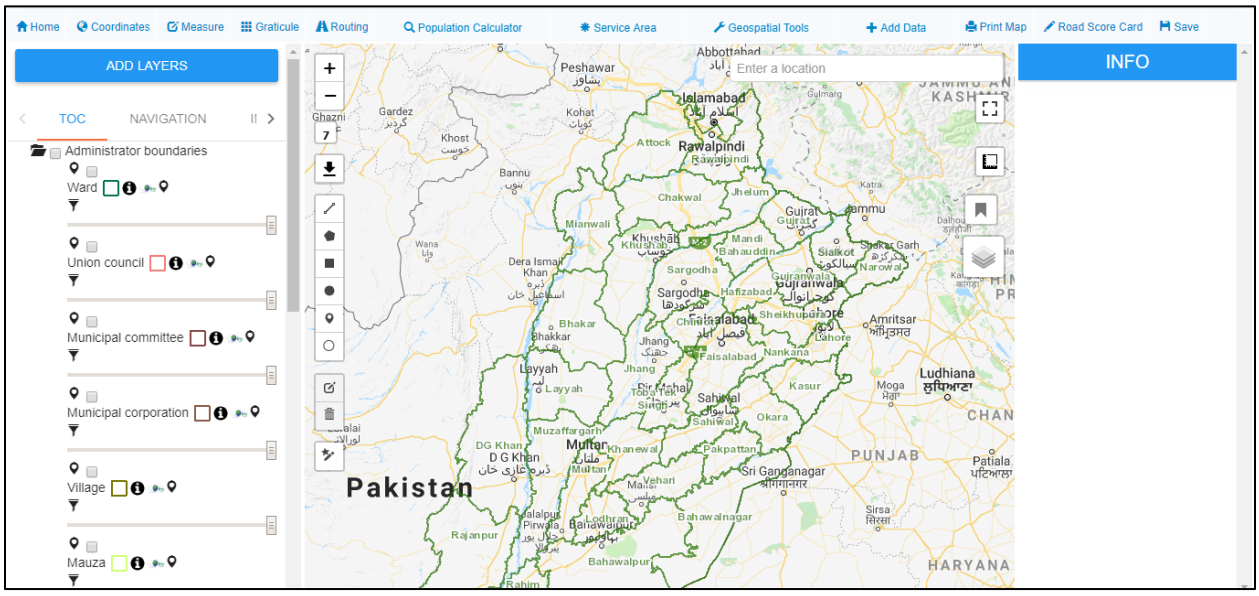
While understanding urban challenges as wicked problems that rarely, if ever, have ‘solutions’, and recognizing that there is always a political element, there is a need to investigate big data and smart cities to better discern their likely implications. In Pakistan, urban planning faces a number of serious challenges, of which the use of data is only one.

Appendix: Project Profiles

This section presents three examples of initiatives to increase the use of data for planning in Punjab. The examples are as follows:

1. IRIS Planner, a planning support system developed by the Urban Unit.
2. Eyris, a GIS-enabled decision support system developed by the Urban Unit.
3. Punjab Safe City Authority, a monitoring and surveillance system that represents the cutting-edge of installed government technology infrastructure in Pakistan.

Profile: IRIS Planner



Screenshot of IRIS Planner homepage

Overview

The GIS team of the Urban Unit built a planning support system for urban planners at the Urban Unit and urban policymakers at the Planning & Development Department in Lahore. The system consists of an online interface that integrates spatial datasets from across the province of Punjab and allows users to filter dozens of layers pertaining to different indicators.

Description

The IRIS Planner helps planners analyze an area using criteria that they can select. It provides access to datasets and spatial and analytical tools to users, and allows them to easily visualize, save and export results.

For example, if someone wants to build a new school in a given area, IRIS Planner can identify if there are any other schools in the area, what their salient characteristics are, what the population

density of the area is, how many children live in the area, and what the road network in the area is. Using this information, planners can pinpoint locations that would be suitable, addressing the most critical need and providing the best access to the largest number of people.

An exciting feature is the ability to get data about a region by drawing a shape on a map. A planner can draw a polygon or a circle on a map and they would be able to see relevant indicators for the selected area.

In the words of a senior public servant, the IRIS Planner is an unprecedented innovation in Punjab and will completely transform the planning process. Most proposals that come through departments, according to him, represent nothing more than political wishes and are not based on any analysis or arguments. This tool is a way to improve governance in the province.

Government Agency: The IRIS Planner was developed by the Urban Unit's GIS and Punjab Spatial Strategy teams and is included as a module of the Eyris product that the Urban Unit is developing. It is intended to be used by planners working at the Urban Unit as well as other departments, delivered through the provincial Planning & Development department and its attached department, the Directorate General of Monitoring & Evaluation.

Benefits:

The most direct benefit of this tool is that it makes information easily available to planners. As the tool is linked to the Urban Unit database, the data on the tool can be updated in real time, as new data comes in. In addition, data can be exported in a number of GIS-compatible formats.

The other major benefit is in changing the process. IRIS Planner can encourage better quality proposals as well as improve both the speed as well as the rigor of the evaluation process. This

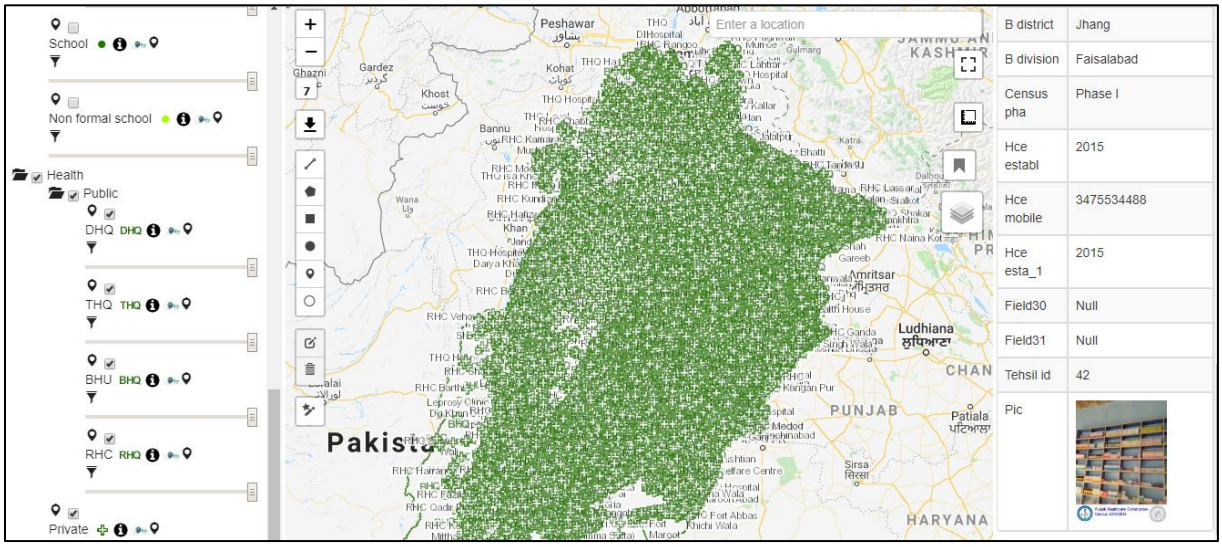
should lead to better projects being funded and greater impact over time. The tool also makes it harder for pet political projects to get through the pipeline and may change the norms of the planning process in Punjab.

Challenges:

A major test that the IRIS Planner faces is whether it will be used or not. It is not clear what mechanisms are in place to enforce a new process. This is particularly relevant in an election year, as Punjab may see a new government which could risk discarding change management initiatives put in place by the previous government.

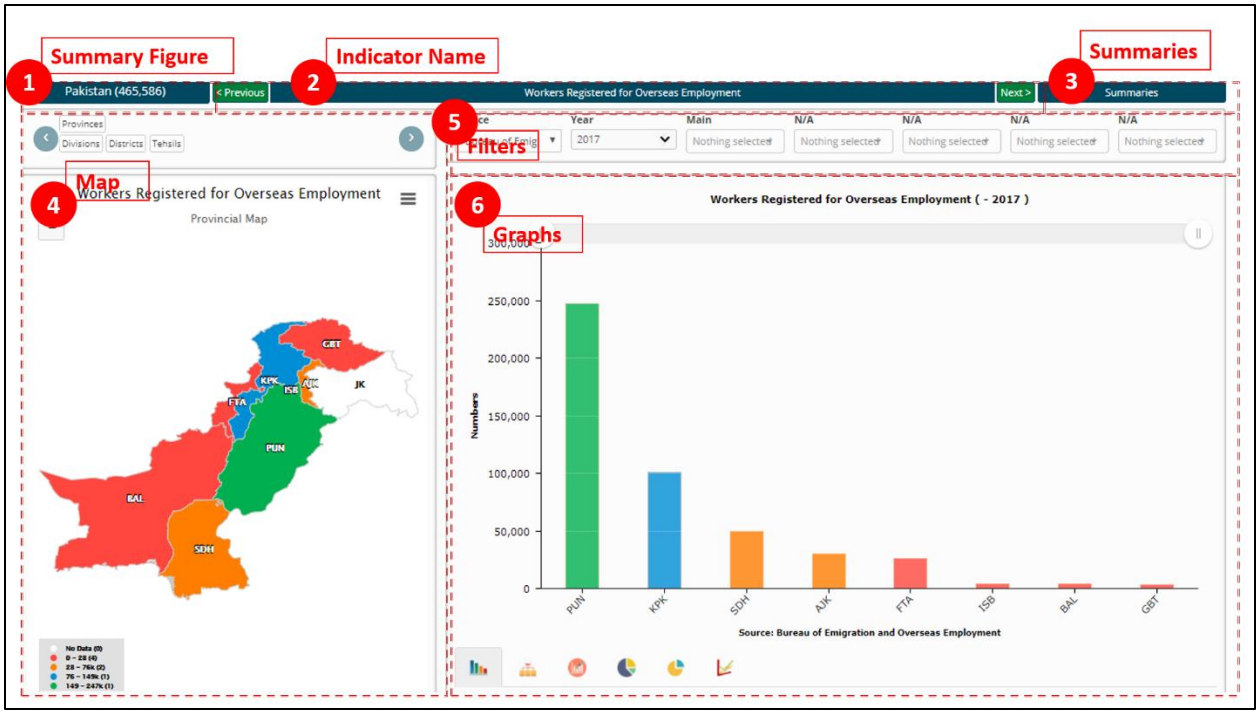
Another challenge is that the tool can only work with the data that it has been provided with. Data quality issues could undermine the value of its use, and so it is critical to ensure that data is vetted before being entered into the database.

Finally, the IRIS Planner has not actually been used yet, and it is not clear whether it has gone through meaningful user testing. One planner I spoke to said that they had only received a single demonstration of the tool but did not feel that it was ready to be used yet.



Screenshot of IRIS Planner showing health facilities in Punjab, with one selection.

Profile: Eyris



Screenshot of Eyris, showing different information displayed on a single screen.

Overview

The Urban Unit has built an information management and decision support system for the planners and policymakers in Pakistan. The system, known Eyris, consists of an online interface that aggregates information across all government ministries and departments. Eyris can be used to track indicators and development schemes, share and retrieve data, and support spatial analysis with the embedded IRIS Planner tool (described separately)

Description

Eyris represents a significant effort to centralize and simplify access to information across the government of Pakistan. In its current iteration it serves the Punjab government, which consists

of 29 sectors and numerous other attached departments that handle the affairs of Pakistan's second-largest and most populous province with over 110 million people.

The system gathers a variety of types of data, including data tables, spatial data, key documents such as meeting minutes to track key decisions, and all project proposals (known as PC-1s). The goal is to upload all data from the year 2000 onwards. This data is carefully extracted, structured and uploaded such that it is tagged and cross-referenced according to geography, sector (i.e. health, education), unit (referring to specific departments) and other relevant information. This functionality makes it a valuable tool for planning, analysis, monitoring & evaluation.

For example, it is now possible to track budgetary allocations to different sectors. A user can see utilization of funds and compare the performance of different departments and lower-tier governments.

Eyris also addresses the issue of data provision and sharing. Whereas in the past it was difficult for public servants to find the information they were looking for, now it is accessible in a few clicks. In addition, the system also provides a user-friendly mechanism for sharing files.

The underlying goal of this "Google of Punjab" is to enable coordination and evidence-based decision making across government. It is an important component of a larger change management initiative.

Government Agency:

Eyris has been developed by the Urban Unit, with the GIS and IT teams working together in close coordination. It is in the process of being completed and has been presented to the provincial Planning & Development Department already. The goal is to make it accessible to all

departments and shift responsibility for uploading data to them. The Urban Unit's role, after development, is to facilitate implementation by providing training and troubleshooting.

Benefits:

Eyris's benefits include making evidence accessible to all stakeholders, enabling easy access and search for data, providing a structured data repository and data-driven summaries, and creating geographical and graphical representation of data and development schemes. It also serves to standardize nomenclature across government and encourage data-driven evaluation of proposals.

Beyond these, a key benefit is to build the capacity of government departments and change the culture to one where data and information sharing are seamlessly integrated instead of being exceptional. It will begin by training over 50 staff at the Planning & Development Department. These staff, and others across the government, will all develop a 'feel' for data through their constant interaction with it.

Challenges:

This project faces the classic chicken-and-egg problem of technology readiness and the capacity to use that technology appropriately. There is a concerted effort to ensure that training is delivered to ensure that Eyris is used, but there are a number of factors that could undermine usage.

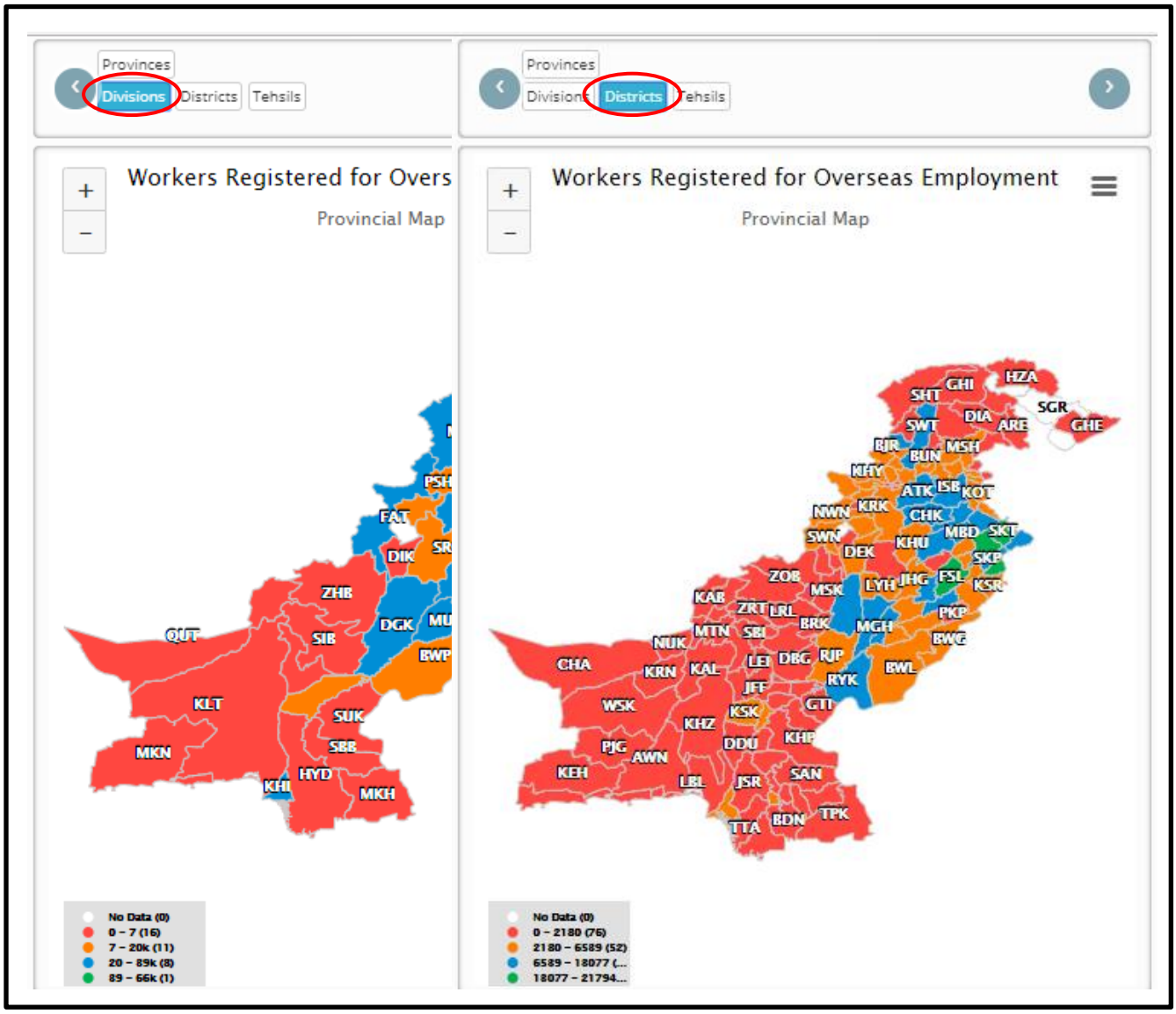
Not everyone will be able to understand how to use the tool because of limited technological literacy. They may also not need, or be interested in, the full range of capabilities that the tool provides (just as few people use all of the functions on their smartphones). Even if they are fully trained and need to use it, they may not be motivated to use it regularly, or may not integrate it

into their workflow. As a result, ongoing engagement and support will be key to enable Eyris to reach its potential.

Another set of challenges revolves around the data itself. This data is still being added, and may be incomplete, missing or incorrect. In fact, the data uploaded to this tool will suffer from the same limitations that all other data in Pakistan face. For example, district level data for a number of indicators will still be missing. Eyris does not address issues related to data collection.

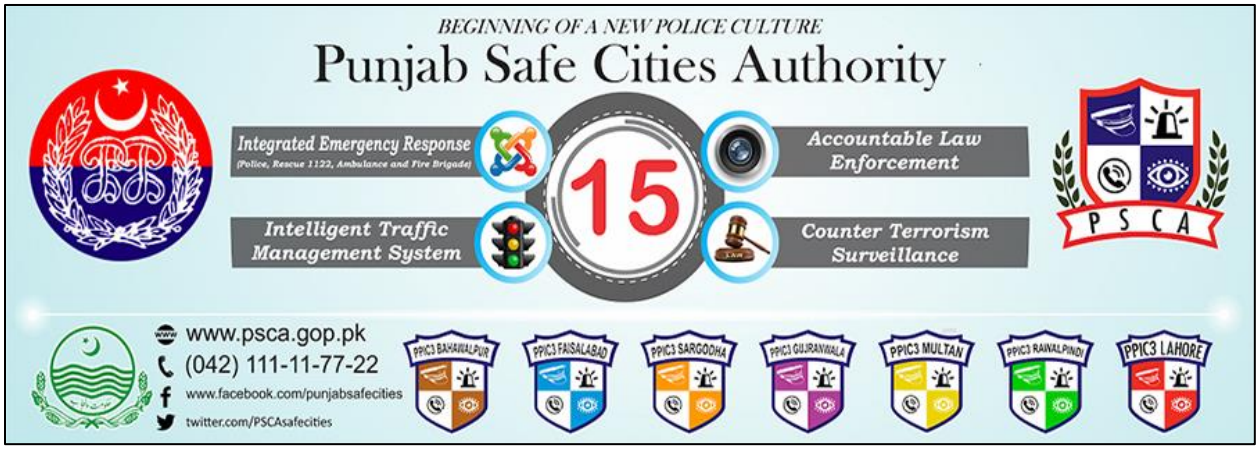
Components:

1. Provincial Level Indicator Database
2. Provincial Level Spatial Database
3. Data Indexing & Searching
4. Indicators Analysis & Presentation
5. PC1 Impact Evaluator
6. IRIS Planner



Screenshot of Eyris – Showing data for an indicator at two geographic scales.

Profile: Punjab Safe Cities Authority



Punjab Safe Cities Authority homepage slider

Overview

The Punjab Safe Cities Authority (PSCA) has built a monitoring and surveillance system for Punjab Police in Lahore. The system consists of thousands of cameras and other sensors installed in over 1,500 locations across Lahore and is headquartered at the Punjab Police Integrated Command, Control and Communication Center (PPIC3). It serves two major purposes: to help the police carry out law enforcement, and to centralize emergency and rescue dispatch services such as ambulances, firefighting, and police.

Description

This project is arguably the most advanced assembly of technology infrastructure being used by a government agency in Pakistan. It is also one of the few examples of government collection and use of big data. Approximately 25 terabytes of various types of data are collected every month through an array of sensors, cameras, microphones, and mobile applications.

The Punjab Police uses this information in a variety of functions, including detecting crimes, carrying out counter-terrorism surveillance, tracking and monitoring vehicles and people, processing First Information Reports (FIRs), handling the city-wide emergency helpline (the local equivalent of 911), and displaying messages on electronic signs across the city. They also use it to help solve criminal cases, such as helping find stolen vehicles and resolving kidnappings.

One of the most well-known features of the system is the e-challan, which allows tickets for traffic violations to be issued. Cameras can detect violations such as illegally crossing signals, identify the license plate of the vehicle and issue a ticket to be sent to the postal address of the registered owner of the car. The driver can then pay via their bank or through their mobile phone.

Beyond improving police intelligence and emergency response times, the Safe Cities initiative is explicitly a way to transform policing in Punjab by supporting evidence-based decision making and coordination between different agencies. The PSCA slogan is "*Beginning of a new police culture*", and a concept paper from 2015 explicitly describes a change management process. The PSCA IT team also spoke about developing indigenous technical expertise and ending vendor dependency.

Government Agency: The Lahore PPIC3 is managed by the Punjab Safe Cities Authority, an autonomous body created in 2015 specifically for this purpose. The PPIC3 was developed by Chinese telecom giant Huawei, who won a contract for Rs. 12 billion (approximately USD $100 million), with consulting and technical support provided by NESPAK and Arup. It is currently operated by the Punjab Police and is located in the police headquarters at Qurban Lines in the

heart of Lahore. PSCA provides ongoing technical expertise and handles maintenance and upgrading.

The Lahore PPIC3 was developed on the heels of another similar but smaller project in Islamabad, which was developed by the federal government for the National Database and Registration Authority between 2009-2016. The Lahore initiative is part of the broader Punjab Safe City program and was launched in 2015, completing in 2018. The PPIC3 is currently being expanded to six other cities in Punjab: Rawalpindi, Multan, Faisalabad, Gujranwala, Bahawalpur, and Sargodha.

Benefits:

PSCA lists the following benefits: improved service delivery (integrated emergency services); traffic management; monitoring law and order; electronic evidence collection; crime reduction; secure police communication, and accountability.

Another set of benefits revolves around the public sector transformation that this project represents. There is significant capacity development taking place, both in terms of institutional knowledge about how to implement large scale technology infrastructure and collaborate more effectively, as well as attracting and training skilled human resources. The Lahore PPIC3 employs approximately 800 people, most under the age of 35 and with degrees in computer science, engineering or information technology. All of them undergo a month of training to operate advanced tools and are paid competitive salaries.

Challenges:

The primary challenge identified by PSCA representatives was that this project entailed privacy violations. They recognized this as a serious issue but also stated that there are no laws to protect the privacy of citizens in Pakistan.

The lack of data sharing and analysis is another key challenge. At present, most of the data collected is discarded after a month. It was not clear if there were any defined processes to analyze the data; most success stories revolve around detecting one-off incidences instead of identifying patterns. Very little of the data is shared outside the PSCA, although in principle the PSCA is open to sharing its data.

Another set of issues revolves around operational challenges. Approximately three-fourths of the 387,000 calls received to the helpline were classified as hoax calls, for example. Asset loss is a significant concern as hardware such as cameras, batteries and fiber-optic cable are often stolen or destroyed.

Finally, PPIC3 faces a number of growth challenges. The scope and ambition of a transformation project such as this one necessarily means that it takes significant time and concerted effort to yield long-term results, although the ability to demonstrate quick wins such as number of recovered kidnapped persons is helpful. In addition, there is a lack of skilled human resources, especially for senior positions. Finally, the sheer costs of technology acquisition and deployment could curtail growth.

Components:

1. Over 8,000 cameras
2. Over 300 panic buttons
3. Over 1,500 locations
4. 1,700 km long dedicated fiber line
5. 168 sites with monitored traffic signals
6. Two mobile applications, one for public safety and one for the security of women.

The Operation Monitoring Centre of PPIC3 Lahore

References

Adam Smith International. (2018). Supporting evidence-based policymaking in Pakistan Retrieved from https://www.adamsmithinternational.com/case-study/supporting-evidence-based-policymaking-in-pakistan

Ahmed, A., & Ahsan, H. (2011). Contribution of Services Sector in the Economy of Pakistan (pp. 1-18, Working paper No. 79). Islamabad: Pakistan Institute of Development Economics.

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

Ahmed, Z. (2017). E-Challan System Gets Operational in Lahore. Pakwired. Retrieved from https://pakwired.com/e-challan-system-gets-operational-in-lahore/

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

Alam, K. (2012). Bucking trends, Teradata invests in Pakistan. Retrieved July 31, 2018, from https://tribune.com.pk/story/484291/bucking-trends-teradata-invests-in-pakistan/

Ali, R. (2013). Estimating Urbanization. The Urban Gazette, (December).

Ali, S., & Rumi, R. (2013). Implementing Urbanisation in Pakistan. In K. Bajwa, Urban Pakistan: Frames for Reading and Imagining Urbanism. Oxford University Press.

Ali, W. (2018). Here's What Pakistan's Consumer Data Protection Framework Should Look Like. ProPakistani. Retrieved from https://propakistani.pk/2018/05/08/heres-what-pakistans-consumer-data-protection-framework-should-look-like/

ALMEC Corporation - Project Summary ::. Retrieved from http://almec.org/projects/D04.html

Andreessen Horowitz. (2017). Companies, Networks, Crowds [Podcast]. Retrieved from https://soundcloud.com/a16z/machine-platform-crowd

Anwar, N. H. (2012). Thinking Beyond 'Engines of Growth': Re-Conceptualizing Urban Planning Discourse in Pakistan, 35(March), 223–238.

Anwar, N., & Ahmed, N. (2017). Census realities. DAWN. Retrieved from https://www.DAWN.com/news/1357320/census-realities

Arby, M., Malik, M., & Hanif, M. (2010). The size of informal economy in Pakistan. Pakistan: Munich Personal ReREc Archive.

Attaa, A. (2018). Govt Officials in Punjab Allegedly Sold NADRA Records, Call and Police Data of Pakistani Citizens. ProPakistani. Retrieved from https://propakistani.pk/2018/05/07/govt-officials-in-punjab-allegedly-sold-nadra-records-call-and-police-data-of-every-pakistani/

Bao, J., He, T., Ruan, S., Li, Y., & Zheng, Y. (2017). Planning Bike Lanes based on Sharing-Bikes' Trajectories. Retrieved from https://www.microsoft.com/en-us/research/publication/planning-bike-lanes-based-sharing-bikes-trajectories/

Batty, M. (1994). A Chronicle of Scientific Planning: The Anglo-American Modeling Experience. Journal of the American Planning Association, 60(1), 7–16.

Batty, M. (1997). The computable city. International Planning Studies, 2(2), 155–173.

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

Behn, R. (2007). What All Mayors Would Like to Know About Baltimore's CitiStat Performance Strategy. IBM. Retrieved from http://web.pdx.edu/~stipakb/download/PerfMeasures/CitiStatPerformanceStrategy.pdf

Bettencourt, L. M. A. (2014). The Uses of Big Data in Cities. Big Data, 2(1), 12–22.

Bhatti, Z. K., Kusek, J. Z., & Verheijen, T. (2015). Logged on: smart government solutions from South Asia.

Bibri, S. E., & Krogstie, J. (2017). ICT of the new wave of computing for sustainable urban forms: Their big data and context-aware augmented typologies and design concepts. Sustainable Cities and Society, 32(7030), 449–474.

boyd, danah, & Crawford, K. (2012). Critical questions for big data. Information, Communication & Society, 15(5), 662–679.

Burton, I. (1963). the Quantitative Revolution and Theoretical Geography. Canadian Geographer, 7(4), 151–162.

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. Retrieved from 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

Campbell, D. (1979). Assessing the impact of planned social change. Evaluation and Program Planning, 2(1), 67-90.

Campbell, H. (1995). A social interactionist perspective on computer implementation. Journal of the American Planning Association, (62) 99–107.

Canadian Institute of Planners (2017). Annual Report [Ebook] Retrieved from http://www.cip-icu.ca/Files/Annual-Report/22673-CIP-Annual-Report-v2.aspx

Caprotti, F., Cowley, R., Datta, A., Broto, V. C., Gao, E., Georgeson, L., … Joss, S. (2017). The New Urban Agenda: key opportunities and challenges for policy and practice. Urban Research & Practice, 00(00), 1–12.

Caro, R. (1974). The Power Broker: Robert Moses and the Fall of New York. New York City: Knopf.

Chun, S. A., Shulman, S., Sandoval, R., & Hovy, E. (2010). Government 2.0: Making connections between citizens, data and government. Information Polity, 15(1–2), 1–9.

Cisco. (2018). Cisco Global Cloud Index: Forecast and Methodology, 2016–2021 (pp. 23-24). San Jose: Cisco Systems Inc. Retrieved from https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.pdf

CityDashboard London. (2018.) Retrieved July 31, 2018, from http://citydashboard.org/london/

Data Driven Decision Making | Information Technology University. (2018). Retrieved from https://itu.edu.pk/it-capacity-building/data-driven-decision-making/

Dalton, C. M., & Thatcher, J. (2014). What does a critical data studies look like?

Datta, A. (2015). New urban utopias of postcolonial India. Dialogues in Human Geography, 5(1), 3–22.

Davoudi, S. (2006). Evidence-Based Planning. DisP - The Planning Review, 165(2), 14–24.

DAWN. (2005). ISLAMABAD: Unplanned urbanization causing negative factors: report. Retrieved from https://www.DAWN.com/news/380124

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.

Desouza, K. C., & Jacob, B. (2017). Big Data in the Public Sector: Lessons for Practitioners and Scholars. Administration & Society, 49(7), 1043–1064.

Diebold, F. X. (2012). A Personal Perspective on the Origin(s) and Development of "Big Data." PIER Working Paper No. 13-003.

Diamantoulakis, P. D., Kapinas, V. M., & Karagiannidis, G. K. (2015). Big Data Analytics for Dynamic Energy Management in Smart Grids. Big Data Research, 2(3), 94-101.

Domínguez, G. (2014). Pakistan's urbanization: 'A challenge of great proportions'. DW. Retrieved from https://p.dw.com/p/1EDDP

Eubanks, V. (2017). Automating inequality. New York: St. Martin's Press.

Faludi, A., & Waterhout, B. (2006). Introducing Evidence-Based Planning. DisP - The Planning Review, 42(165), 4–13.

Finance Department. (2017). Revision of Basic Pay Scales & Allowances of Civil Servants of the Punjab Government (pp. 1-6). Lahore: Government of Punjab.

Finch, G., Davidson, S., Kirschniak, C., Weikersheimer, M., Reese, C., & Shockley, R. (2014). Analytics: The Speed of Advantage - Why data-driven organizations are winning the race in today's marketplace (p. 28). IBM Global Business Services.

Finch, G., Davidson, S., Kirschniak, C., Weikersheimer, M., Rodenbeck Reese, C., & Schockley, R. (2014). Analytics: The speed advantage. Executive Report, 23.

Florescu, D., Karlberg, M., Reis, F., Rey, P., Castillo, D., Skaliotis, M., & Wirthmann, A. (2014). Will 'big data' transform official statistics?

Flyvbjerg, B., & Richardson, T. (2002). Planning and Foucault: In Search of the Dark Side of Planning Theory. Planning Futures: New Directions for Planning Theory, 44–63.

Foth, M. (2009). Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City (Vol. 75).

French, S. Barchers, C. Zhang, W. (2015). Moving beyond operations: leveraging Big Data for urban planning decisions. 14th International Conference on Computers in Urban Planning and Urban Management, 194-1–16.

Future Cities Catapult. (2017). Smart City Strategies: A Global Review.

Gallup Pakistan. (2018). Big Data Analysis Reports. Retrieved from http://gallup.com.pk/polls/gallup-history-project/big-data-analysis/

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. International Journal of Information Management, 35(2), 137–144.

Gazette of Pakistan (2010). President's Order No. 5 of 2010. Islamabad: Government of Pakistan, pp. http://finance.gov.pk/nfc/presidential_order_5_2010.pdf.

Geertman, S. (2017). PSS: Beyond the implementation gap. Transportation Research Part A: Policy and Practice, 104, 70–76.

Geertman, S., & Stillwell, J. (2004). Planning support systems: An inventory of current practice. Computers, Environment and Urban Systems, 28(4), 291–310.

Geo News. (2017). Pakistan to become first country to test 5G: Anusha Rahman. Retrieved from https://www.geo.tv/latest/147694-pakistan-to-become-first-country-to-test-5g-anusha-rahman

Government of Pakistan. (2016). Year Book (p. 73). Islamabad: Pakistan Bureau of Statistics.

Greenspan, E. (2016). Top-Down, Bottom-Up Urban Design. New Yorker. Retrieved from https://www.newyorker.com/business/currency/top-down-bottom-up-urban-design

Haider, M., & Badami, M. G. (2010). Urbanization and Local Governance Challenges in Pakistan. Environment and Urbanization ASIA, 1(1), 81–96.

Haider, M. (2014). Pakistan’s Urbanization Challenges: Transport and Mobility. In M. Kugelman, Pakistan's Runaway Urbanization: What Can Be Done?. Washington D.C.: Wilson Center.

Hall, P. (2014). Cities of Tomorrow: An Intellectual History of Urban Planning and Design Since 1880 (Vol. 13).

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

Hasan, A. (2012). The impending migration. DAWN. Retrieved from https://www.DAWN.com/news/768650

Hasan, A. (2012b). Planning and Its Assumptions. Presentation, Karachi, Pakistan.

Hasan, A., & Raza, M. (2013). Migration and small towns in Pakistan. Karachi: Oxford University Press.

Hashem, I. A. T., Chang, V., Anuar, N. B., Adewole, K., Yaqoob, I., Gani, A., … Chiroma, H. (2016). The role of big data in smart city. International Journal of Information Management, 36(5), 748–758.

Hoang, M., Zheng, Y., & Singh, A. (2016). FCCF: Forecasting Citywide Crowd Flows Based on Big Data. Retrieved from https://www.microsoft.com/en-us/research/publication/forecasting-citywide-crowd-flows-based-big-data/

Hollands, R. G. (2015). Critical interventions into the corporate smart city. Cambridge Journal of Regions, Economy and Society, 8(1), 61–77.

Hull, M. (2012). Government of Paper: The Materiality of Bureaucracy in Urban Pakistan. Berkeley: University of California Press

IBM Watson Marketing. (2016). 10 Key Marketing Trends for 2017 and Ideas for Exceeding Customer Expectations [PDF] (p. 3). Retrieved from https://public.dhe.ibm.com/common/ssi/ecm/wr/en/wrl12345usen/watson-customer-engagement-watson-marketing-wr-other-papers-and-reports-wrl12345usen-20170719.pdf

IDC. (2014). Executive Summary: Data Growth, Business Opportunities, and the IT Imperatives. Retrieved from https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm

IDC. (2017). Big Data and Business Analytics Revenues Forecast to Reach $150.8 Billion This Year, Led by Banking and Manufacturing Investments, According to IDC. (2017). Retrieved from https://www.idc.com/getdoc.jsp?containerId=prUS42371417

Iliadis, A., & Russo, F. (2016). Critical data studies: An introduction. Big Data & Society, (December), 1–7.

Institutional Reform Group. (2012). Showcasing Service Delivery Innovations in Punjab - Punjab Governance Forum 2012 [Ebook]. Retrieved from https://irg.punjab.gov.pk/sites/irg.pitb.gov.pk/files/Showcasing_Service_Delivery_Innovations_Punjab.pdf

InnovatePk. (2018). Pakistan Enters the Age of Big Data with PKR1.53 Billion National Center. Retrieved July 31, 2018, from http://innovatepk.org/2018/05/18/pakistan-enters-the-age-of-big-data-with-pkr153b-national-center/

Iqbal, A. (2014). Urbanization in Pakistan: A Perspective from Government. In M. Kugelman, Pakistan's Runaway Urbanization: What Can Be Done?. Washington D.C.: Wilson Center.

Jamali, H. (2018). Mathew S. Hull, Government of Paper: The Materiality of Bureaucracy in Urban Pakistan. Retrieved from https://journals.openedition.org/samaj/3969

Jan, B., & Iqbal, M. (2008). Urbanization Trend and Urban Population Projections of Pakistan Using Weighted Approach. Sarhad Journal of Agriculture, 24(1). Retrieved from http://www.aup.edu.pk/sj_pdf/Urbanization%20trend%20and%20rban%20population%20projections%20of%20pakis.DOC.pdf

Khan, A. (2017). Trading data. Retrieved July 31, 2018, from https://www.thenews.com.pk/magazine/money-matters/194614-Trading-data

Kitchin, R. (2014a). The real-time city? Big data and smart urbanism, (November 2013), 1–14.

Kitchin, R. (2014b). The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. London: Sage Publications

Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. Big Data & Society, 3(1), 205395171663113.

Klopp, J. M., & Petretta, D. L. (2017). The urban sustainable development goal: Indicators, complexity and the politics of measuring cities. Cities, 63, 92–97.

Klosterman, R. E. (2013). Lessons Learned About Planning. Journal of the American Planning Association, 79(2), 161–169.

Komninos, N. (2009). Intelligent cities: towards interactive and global innovation environments. International Journal of Innovation and Regional Development, 1(4), 337.

Krivý, M. (2018). Towards a critique of cybernetic urbanism: The smart city and the society of control. Planning Theory, 17(1), 8–30.

Kugelman, M. (2014). Pakistan's Runaway Urbanization: What Can Be Done? Washington D.C.: Wilson Center.

Kwan, M. P., & Schwanen, T. (2009). Quantitative revolution 2: The critical (re)turn. Professional Geographer, 61(3), 283–291.

Klosterman, R. (2008). A New Tool for a New Planning: The What if? Planning Support System. In R. Brail, Planning Support Systems for Cities and Regions. Cambridge: Lincoln Institute of Land Policy.

Kunicova, J., & Bhatti, Z. (2015). Building trust in the government one text at a time. The World Bank - Blog. Retrieved from https://blogs.worldbank.org/governance/endpovertyinsouthasia/governance/ppps/building-trust-in-government-though-mobile-messaging

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The Parable of Google Flu: Traps in Big Data Analysis. Science, 343(6176), 1203-1205.

Leadbeater, C. (2011). The Civic Long Tail. London: Demos. Retrieved from http://www.demos.co.uk/files/Civic_long_tail_-_web.pdf?1315915449%5Cnhttp://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+Civic+Long+Tail#1

LUMS. (2017). MBA Programme. [Brochure] Lahore, Pakistan. Retrieved from https://lums.edu.pk/sites/default/files/research-publication/mba_brochure_25_august_1.pdf

Madera, C. (2010). The Future of Cities in the Internet Era. Retrieved from https://nextcity.org/daily/entry/the-future-of-cities-in-the-internet-era

Malik, T. (2015). Big Data in Government's Service to Citizens and State. Retrieved July 31, 2018, from https://www.forbes.com/sites/teradata/2015/01/30/big-data-in-governments-service-to-citizens-and-state/

Marcus, G., & Davis, E. (2013). What Nate Silver Gets Wrong. Retrieved from https://www.newyorker.com/books/page-turner/what-nate-silver-gets-wrong

Marshall, A., Mueck, S., & Shockley, R. (2015). How leading organizations use big data and analytics to innovate. Strategy and Leadership, 43(5), 32–39.

Mayer-Schoenberger, Viktor Cukier, K. (2013). The Rise of Big Data. Foreign Affairs.

Miller, C. A. (2005). New civic epistemologies of quantification: Making sense of indicators of local and global sustainability. Science Technology and Human Values, 30(3), 403–432.

Martine, G., Eustaquio Alves, J., & Cavenaghi, S. (2013). Urbanization and fertility decline: Cashing in on structural change (pp. 1-44). London, UK: International Institute for Environment and Development. Retrieved from http://pubs.iied.org/pdfs/10653IIED.pdf

Matsakis, L. (2018). Your Smartphone Choice Could Determine If You Get a Loan. Wired. Retrieved from https://www.wired.com/story/your-smartphone-could-decide-whether-youll-get-a-loan

Mehboob, A. (2018). Give LGs a chance. DAWN. Retrieved from https://www.DAWN.com/news/1304711

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

Ministry of IT & Telecom. (2018a). Digital Pakistan Policy. Retrieved from http://moitt.gov.pk/userfiles1/file/DIGITAL%20PAKISTAN%20POLICY.pdf

Ministry of Information Technology & Telecommunication. (2018b). PERSONAL DATA PROTECTION BILL 2018. Government of Pakistan

MIT Technology Review Pakistan. (2018). To boost Big Data analytics, Pakistan now has a national center with 12 affiliated laboratories. Retrieved from http://www.technologyreview.pk/to-boost-big-data-analytics-pakistan-now-has-a-national-center-with-12-affiliated-laboratories/

Mustafa, U. (2011). Fiscal federalism in Pakistan: The 7th National Finance Commission award and its implications (No. 73). Islamabad: Pakistan Institute of Development Economics. Retrieved from https://core.ac.uk/download/pdf/6489620.pdf

Nabi, S. (2015). Consortium for Development Policy Research Urban Development in Punjab: A Political Economy Analysis Consortium for Development Policy Research. Consortium for Development Policy Research.

Nam, T., & Pardo, T. A. (2011). Conceptualizing Smart City with Dimensions of Technology, People, and Institutions. In 12th Annual International Conference on Digital Government Research Conceptualizing (pp. 282–291).

NADRA Pakistan. (2018a). International Projects. Retrieved from https://www.nadra.gov.pk/international-projects/

NADRA Pakistan. (2018b). Safe City. Retrieved from https://www.nadra.gov.pk/solutions/security-solutions/safe-city-solutions/

Naviwala, N. (2018). Punjab: regime of fear. Retrieved from https://www.dawn.com/news/1270283

Nott, C. (2015). A maturity model for big data and analytics. Retrieved from https://www.ibmbigdatahub.com/blog/maturity-model-big-data-and-analytics

O'Donnell, M. A. (2013). on shenzhen speed. Retrieved July 31, 2018, from https://shenzhennoted.com/2013/04/27/on-shenzhen-speed/

Pakistan Bureau of Statistics. (2018a). Data Dissemination. Retrieved from http://www.pbs.gov.pk/content/data-dissemination

Pakistan Bureau of Statistics. (2018b). Methodology. Retrieved from http://www.pbs.gov.pk/content/methodology-1

Pakistan Council of Architects and Town Planners. (2018). Pakistan Council of Architects and Town Planners Handbook [Ebook]. Retrieved from http://www.pcatp.org.pk/images/pdf/PCATP%20Hand%20Book.pdf

Pasha, H. A. (2018). Growth and Inequality in Pakistan. Friedrich Ebert Stiftung.

Pelzer, P., Geertman, S., Van der Heijden, R., & Rouwette, E. (2014). The added value of Planning Support Systems: A practitioner's perspective. Computers Environment and Urban Systems2, 48, 16–27.

Pakistan Telecommunication Authority. (2017). PTA Annual Report 2017 [Ebook]. Islamabad. Retrieved from https://www.pta.gov.pk/assets/media/ann_rep_2017.pdf

Pettit, C., Klosterman, R., Nino-Ruiz, M., Widjaja, I., Russo, P., & Tomko, M. et al. (2013). The Online What if? Planning Support System. Lecture Notes in Geoinformation And Cartography, 349-362.

Palvia, S., & Sharma, S. (2018). E-Government and E-Governance: Definitions/Domain Framework and Status around the World [Ebook]. Retrieved from http://csi-sigegov.orgwww.csi-sigegov.org/1/1_369.pdf

Poon, L. (2017). Finally, Uber Releases Data to Help Cities With Transit Planning. Retrieved from https://www.citylab.com/transportation/2017/01/finally-uber-releases-data-to-help-cities-with-transit-planning/512720/

Press, G. (2016). Cleaning Big Data: Most Time-Consuming, Least Enjoyable Data Science Task, Survey Says. Retrieved from https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/

Punjab Bureau of Statistics. (2018). Multiple Indicator Cluster Survey (MICS). Retrieved from http://www.bos.gop.pk/mics

Punjab IT Board. (2017). Digital Punjab: Enhancing Public Services Through Technology. Retrieved from http://booklet.pitb.gov.pk/mobile/Flipbook.aspx

Punjab Safe Cities Authority. (2018). PSCA RELEASES IC3 PERFORMANCE REVIEW FOR MAY 2018. Retrieved from http://psca.gop.pk/PSCA/2018/06/08/84859/

Qadeer, M. A. (1996). An assessment of Pakistan's urban policies, 1947-1997. Pakistan Development Review, 35(4 PART 2), 443–465.

Qadeer, M. (1999). 'Urbanisation of Everybody', Institutional Imperatives, and Social Transformation in Pakistan. The Pakistan Development Review, 38(4 Part 2), 1193–1210.

Qadeer, M. (2014). Do's and Don'ts of Urban Policies in Pakistan. In M. Kugelman, Pakistan's Runaway Urbanization: What Can Be Done?. Washington D.C.: Wilson Center.

Rabari, C., & Storper, M. (2015). The digital skin of cities: urban theory and research in the age of the sensored and metered city, ubiquitous computing and big data, (October 2014), 27–42.

Raja, M. (2016). Unplanned urbanisation: Rawalpindi spiralling out of control. The Express Tribune. Retrieved from https://tribune.com.pk/story/1249910/unplanned-urbanisation-rawalpindi-spiralling-control/

Randhawa, S. (2017). Lahore: a victim of unplanned urbanisation. Pakistan Today. Retrieved from https://www.pakistantoday.com.pk/2017/01/20/lahore-a-victim-of-unplanned-urbanisation/

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. Policy Sciences, 4(2), 155–169.

Rogerson, R. J. (1999). Quality of Life and City Competitiveness. Urban Studies, 36(5), 969–985.

Rogge, N., Agasisti, T., & De Witte, K. (2017). Big data and the measurement of public organizations' performance and efficiency: The state-of-the-art. Public Policy and Administration, 32(4), 263–281.

Roveda, P. (2017). The Online Whatif GIS-based Planning Support System Tool. Retrieved from http://www.cityplanner.it/gistips/the-online-whatif-gis-based-planning-support-system-tool/

Sage, A. (2006). Where's the lane? Self-driving cars confused by shabby U.S. roadways. Reuters. Retrieved from https://www.reuters.com/article/us-autos-autonomous-infrastructure-insig/wheres-the-lane-self-driving-cars-confused-by-shabby-u-s-roadways-idUSKCN0WX131

Saint-Germain, R. (2005). Information Security Management Best Practice Based on ISO/IEC 17799. Information Management Journal. Retrieved from https://pdfs.semanticscholar.org/a94d/b6ae3b0676a975ea094fe818c467350d2a7e.pdf

Sachs, W. (2010). The Development Dictionary (2nd ed.). London: Zed Books.

Saeed, Fahad; Majeed Salik, Kashif; Ishfaq, S. (2016). Climate induced rural-to-urban migration in Pakistan.

Sheehan, H., & Sweeney, S. (2018). The Wire and the World. Jacobin.

Stevens, Hallam; Fearnley, Lyle; Sun, Shirley; Watson, S. (2017). Big Data in Asia: Provocations and Potentials. Big Data & Society.

Schultz, J. (2017). How Much Data is Created on the Internet Each Day? 10.10.2017 by Jeff Schultz [Blog]. Retrieved from https://blog.microfocus.com/how-much-data-is-created-on-the-internet-each-day/

Siddiqui, T. (2014). Pakistan's Urbanization Challenges: Housing for the Low-Income. In M. Kugelman, Pakistan's Runaway Urbanization: What Can Be Done?. Washington D.C.: Wilson Center.

Sigal, I. (2018). These Studies Led to Further Studies. Retrieved from https://placesjournal.org/article/these-studies-led-to-further-studies

Simon, D. (2015). Mr. O'Malley's Bad Math [Web blog post]. Retrieved July 31, 2018, from https://davidsimon.com/omalley-bad-math/

Startup Challenge Summer 2017. (2017). Retrieved from http://www.urbanunit.gov.pk/UU/SCS17

Statistics Division. (2017). Yearbook 2015-16. Islamabad: Ministry of Finance, Economic Affairs, Revenue, Statistics & Privatization.

Strengthening National Statistics in Pakistan. (2016). Retrieved from https://devtracker.dfid.gov.uk/projects/GB-1-108746/documents

Sustainable Development Policy Institute. (2015). Data Gaps and Evidence-based Policy-making in Pakistan. Retrieved from http://sdpi.org/policy_outreach/policy_recommendations_details532.html

Thakuriah, P., Tilahun, N., & Zellner, M. (2017). Seeing Cities Through Big Data. Springer.

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

The Express Tribune. (2011). Government data consolidated under Pakistan Bureau of Statistics. Retrieved from https://tribune.com.pk/story/311187/government-data-consolidated-under-pakistan-bureau-of-statistics/

The Express Tribune. (2014). Déja vu: Fire destroys records at Shaheed-e-Millat, again. Retrieved from https://tribune.com.pk/story/678125/deja-vu-fire-destroys-records-at-shaheed-e-millat-again/

The Gazette of Pakistan. General Statistics (Reorganization Act), 2011 (2018). Islamabad.

The Nation. (2017). Shenzhen Speed is past, Punjab Speed is present. Retrieved from https://nation.com.pk/07-Apr-2017/shenzhen-speed-is-past-punjab-speed-is-present

The Nation. (2018). Smart City: The Hackathon concludes at ITU. Retrieved from https://nation.com.pk/11-Mar-2018/smart-city-the-hackathon-concludes-at-itu

The News. (2017). Population Census 2017: Sindh govt rejects results, mulls APC. The News. Retrieved from https://www.thenews.com.pk/latest/226672-Population-Census-2017-Sindh-govt-rejects-results-mulls-APC

The Urban Unit (2007). Assessment of Land Development and Management Practices in Five Large Cities of Punjab.

The Urban Unit (2018). Punjab Cities Growth Atlas 1995-2015. Retrieved from http://uu.urbanunit.gov.pk/enewsletters/urbanatlas/index.html

The Urban Unit. (2018b). Urban Unit - Code of Conduct.

Townsend, A. (2013). Smart cities. New York City: W.W. Norton.

Townsend, A. (2015). Cities of Data: Examining the New Urban Science. Public Culture, 27(2 76), 201–212.

Trundle, A., Stephan, A., Henderson, H., Kamalipour, H., & Lowe, M. (2016). Habitat III is over, but will its New Urban Agenda transform the world's cities? The Conversation. Retrieved from https://theconversation.com/habitat-iii-is-over-but-will-its-new-urban-agenda-transform-the-worlds-cities-67432

United Nations (1992). The Future is Urban. Retrieved from https://www.un.org/popin/icpd/newslett/92_09/The+Future+is+Urban.html

United Nations Economic and Social Commission for Asia and the Pacific. (2016). State of ICT in Asia and the Pacific 2016: Uncovering the Widening Broadband Divide. Retrieved from https://www.unescap.org/sites/default/files/State%20of%20ICT%20in%20Asia%20and%20the%20Pacific%202016.pdf

United Nations. (2017). New Urban Agenda. Conference on Housing and Sustainable Urban Development (Habitat III).

United Nations (2018). United Nations E-Government Survey 2018 [Ebook]. Retrieved from https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2018-Survey/E-Government%20Survey%202018_FINAL%20for%20web.pdf

UrbanSim. (2018). UrbanSim. Retrieved from http://www.urbansim.com/home

Usman, M. (2013). Karachi's geek squad: Civic Hackathon comes with an app to keep tab on rishwat. The Express Tribune. Retrieved from https://tribune.com.pk/story/532871/karachis-geek-squad-civic-hackathon-comes-with-an-app-to-keep-tab-on-rishwat/

Vazquez, A., & Bhatti, Z. (2013). M-government? – Innovations from Punjab. The World Bank - Blogs. Retrieved from https://blogs.worldbank.org/governance/m-government-innovations-punjab#1

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

Von Solms, B., & Von Solms, R. (2004). The 10 deadly sins of information security management. Computers and Security, 23(5), 371–376.

Vonk, G., Geertman, S., & Schot, P. (2005). Bottlenecks blocking widespread usage of planning support systems. Environment and Planning A, 37(5), 909–924.

What If? (2016). Welcome to What if? - GIS-based planning support system - Explore alternate futures for a community. (2016). Retrieved from http://whatifinc.biz/index.php

Wiig, A. (2015). IBM's smart city as techno-utopian policy mobility. City, 19(2–3), 258–273.

World Bank. (2018) Statistical Capacity score (Overall average) | Data. Retrieved from https://data.worldbank.org/indicator/IQ.SCI.OVRL?locations=PK-XT-XP-XN&view=chart

Yaqoob, I., Hashem, I. A. T., Gani, A., Mokhtar, S., Ahmed, E., Anuar, N. B., & Vasilakos, A. V. (2016). Big data: From beginning to future. International Journal of Information Management, 36(6), 1231–1247.

Yuen, B., & Choi, S. (2012). Making Spatial Change in Pakistan Cities Growth Enhancing.

Yousafzai, W. (2013). 1.6million Fata IDPs living in Peshawar. ReliefWeb. Retrieved from https://reliefweb.int/report/pakistan/16million-fata-idps-living-peshawar

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

Zheng, Y., Capra, L., Wolfson, O., & Yang, H. (2014). Urban Computing. ACM Transactions on Intelligent Systems and Technology, 5(3), 1–55.