POLICY VANGUARD

Policy Vanguard
Vol. 1, No. 1, 2025

RESEARCH ARTICLE

Importance Of Big Data: Pakistan's Struggle With Big Data Governance

Muhammad Umar Farooq Baloch

Research Scholar, Govt And Public policy, National Defence University, Islamabad

Received: May 20, 2025
Accepted: Jun 5, 2025
Published: Jun 26, 2025

Abstract
The digital information landscape and the Internet of Things (IOT) have become the new vitality of society, and the word "data" which is often used interchangeably, has become essential to human existence in the modern world. Artificial intelligence (AI) in data analytics, IoT services, and human action sensors all highlight how important technology is in forming our cyber spectrum. In the age of surveillance capitalism, this study will highlight Pakistan's heavy reliance on data and examine the country's need for a strong big data governance framework. It highlights the absence of a cohesive national framework while pointing out privacy and data management shortcomings. In order to safeguard and use big data for advancements in both economics and security, it also promotes the creation of a central data management system and strict data protection regulations.

Keywords: Artificial intelligence, Big Data, Socio-Economic landscape, cyber spectrum, Data governance

Introduction

Traditionally, state businesses relied on manual data analysis exercises carried out by individuals, which often used to produce inaccurate results and many problems remains unsolved due to limited capacity of human data analysts and lack of modern data vision automation technologies. However, the advent of big data, cloud computing, and now AI have altered the landscape of data analytics completely. Consequently, statecraft has become more sophisticated and data-driven. The use of big data in various stat crafts is expected to continue to evolve and expand in the coming years, as governments and organizations seek to leverage the power of data to make more informed and effective decisions.

Big Data is a term used in analytics involves processing and analysing large and complex data sets using advanced tools and techniques. This allows for the identification of patterns, trends, and insights that would be difficult or

impossible to detect using traditional methods. In the context of statecraft, big data analytics can be used to analyse and interpret data from various sources, including government databases, social media platforms, and sensor networks.

The most prominent attributes of Big Data are referred as 3Vs i.e. **Volume** (Amount of Data), **Variety** (Type or format of Data) and **Velocity** (Data generation/processing speed). It is evident that all 3 Vs demands installation of specialized hardware and software solutions to generate, sort, clean, store, manage, retrieve, share large amount of data as and when required.[i]

Big data plays a crucial role in e-governance by enabling governments to analyse data trends,improve decision- making, and customize services to meet citizen's needs. It acts as a catalyst for trust, efficiency, and innovation within the public sector. By harnessing big data analytics, governments can enhance service delivery, engage with the public more effectively, and streamline governmental processes. This technology expands the horizons of e-governance, making administrative reforms tangible and enhancing the overall governance ecosystem.

Pakistan is 5th largest country in the world in terms of population is having a unique set of socio-economic challenges. Shirking resources and growing populations are impeding national progress and creating new issues like growing clusters of unemployed, under privileged and socio-economically backward masses.[ii] In this context, it is critical for country's economic and social viability to adopt data driven policy frameworks for all the statecraft particularly those related to governance, public services, economy, commerce and trade.

Relevance of Big Data in Pakistan's Socio-Economic landscape

Big data has become increasingly relevant in Pakistan's economic and social landscape, driven by advancements in contemporary technologies such as artificial intelligence (AI), machine learning, and cloud computing. These technologies enable the collection, processing and analysis of vast datasets, which is crucial for informed decision-making in sectors like agriculture, healthcare, and finance. In a country like Pakistan, Adoption of big data offers a way to improve operational efficiency and strategic planning for both public and private sectors where data collection has often been fragmented.Within the economic sphere, big data has the potential to greatly improve the productivity of sectors like banking, retail, and manufacturing. Businesses can streamline operations, cut costs, and boost productivity by evaluating supply chain data, consumer behaviour, and market trends. Real-time data analysis is becoming crucial for innovation in Pakistan's expanding digital economy, which is being driven by the country's increasing use of mobile devices and fintech innovations.[iii] Businesses can better customise products, optimise workflows, and compete in international markets recognitions to the insights gleaned from big data.

Big data has the ability to help Pakistan with some of its most urgent social issues. Predictive analysis, for example, can enhance disease surveillance and management in the healthcare industry, assisting policymakers in addressing public health emergencies like the spread of infectious diseases. Big data in education can be used to better allocate resources and evaluate learning outcomes, resulting in more equitable access to high-quality education. Big data can also be used by the government to plan cities more effectively, especially in developing areas with high infrastructure demands like Karachi and Lahore.

Pakistan has initiated various e-governance initiatives ushering automation and improved services model. Though these initiatives offer a quick and efficient service delivery by various departments to public but at strategic policy level enormous data generated daily by these systems is either wasted or not utilized to improve services, tracking socio-

economically challenged clusters of masses, accountability, crime detection, and transparency.

Overall, the absence of data-driven policies in governance can lead to inefficiencies, inequities, and challenges in addressing the diverse needs of a large population. Implementing data-driven policies is crucial for improving governance, enhancing service delivery, and fostering transparency and accountability in countries like Pakistan. There are number of reasons due to which the importance of data and its exploitation through Big Data, cloud computing, and AI has not fully realized. These challenges range from comprehension to skill gap to privacy issues to out dated system technology in the existing automation and e-governance models.

Pakistan's Big Data landscape

The big data landscape in Pakistan is currently in its beginning stages but is expanding quickly. Some industries are starting to comprehend the potential of big data, while others are trailing behind because of infrastructure and regulatory issues. Big data is being used more and more in industries like retail, telecom, and finance as a result of the rise of digitalisation. However, sectors such as healthcare, education, and agriculture remain significantly underutilized in terms of big data integration. Below is a structured analysis, including a **comparative table** of sector-wise adoption and key initiatives.

Sector	Key Applications	Challenges
Telecom	Customer behavior analytics, Network optimization Personalized marketing	Data privacy concerns
Finance	Fraud detection, Credit wallet analytics (100M+ users)	Lack of unified regulatory framework
Retail	Demand forecasting, Inventory management	Limited digital payment penetration
Healthcare	Paper-based records dominate, Weak real-time disease surveillance	No centralized health data infrastructure
Education	Minimal student performance tracking, Resource allocation gaps	Fragmented data systems
Agriculture	Rare use of IoT/sensor data, No predictive analytics for crops	Poor rural internet connectivity
Government	- NADRA biometric database (100M+ records) - Digital census (2023)	Bureaucratic inefficiencies

Pakistan's financial and telecom sectors are leading the way in leveraging big data for operational efficiency and customer engagement. For example, telecom operators such as Jazz, Telenor, and Zong use data analytics to understand customer patterns, optimize network performance, and design personalized payment campaigns. The state Bank of Pakistan (SBP) has also pushed for digital payment solutions and fintech expansions, which are heavily reliant on data-driven insights. However, in financial sector, banks and fintech companies use big data to assess creditworthiness, manage risks and prevent fraud. The emergence of mobile wallets, with over 100 million users as mentioned above provides massive amounts of transactional data. This data is used to tailor financial products to underserved communities, facilitating financial inclusion. According to Pakistan Telecommunication Authority (PTA), the country's internet user base has crossed 124 million, creating a huge reservoir of data for service optimization.

In realistic converse, Pakistan's healthcare sector is far behind in the adoption of big data. Despite the growing need for prognostic healthcare models, real-time disease surveillance, and improved patient management, the use of data remains largely unstructured. Public health systems frequently keep their records on paper, which makes gathering data ineffective and challenging to analyse. The lack of a comprehensive data infrastructure was revealed through the COVID-19 pandemic by how difficult it was to monitor infection rates in real-time and

distribute resources. Equally, data-informed models have yet to be embraced fully by the education sector. There is a lack of large-scale data on the distribution of resources and performance of students. This lag makes it rigid for decision-makers to correct imbalances and drive educational outcomes.

One of the major issues in Pakistan's big data ecosystem is the weak laws to protect data. The nation is at risk of data privacy and security due to the increasing levels of financial and personal data being collected. The nation must make substantial improvements in the digital infrastructure. Access to high-speed internet is still limited in rural regions, and data centres lag behind in meeting level standards with the rest of the world. The lack of infrastructure prevents efficient storage, processing, and analysis of large data volumes.

In addition, several government-initiated programs aimed at modernizing data gathering, governance, and service delivery have driven the use of big data in Pakistan's key sectors. The digital census and the National Database and Registration Authority (NADRA) are among the most prominent examples. These efforts show the government leveraging big data for better citizen services, policymaking, and governance.

NADRA is an integral part of Pakistan's big data infrastructure. One of the world's largest citizen databases, NADRA has more than 100 million biometric and demographic records. These data are used for social welfare schemes, electoral rolls, security, and identification checks alongside the issuance of Computerised National Identity Cards (CNICs). In order to heighten transparency and heighten the services of the public, NADRA introduced smart cards and real-time data analysis in recent years. For example, the Ehsaas Program, an important social protection program, utilizes NADRA data to ensure targeted support is delivered to those who are eligible for it, thus reducing fraud and inefficiencies. The government can analyse trends and patterns in unemployment, poverty, and demographics through the use of an enormous quantity of data it has collected, enhancing the effectiveness of social safety net programs.

Another significant government initiative using big data is Pakistan's first digital census to be held in 2023, under the guidance of the Pakistan Bureau of Statistics (PBS). Over 240 million individuals are expected to have their information gathered by the digital census, which will help policymakers in resource allocation, infrastructure development, and urban planning. The information will also assist in electoral processes and provide insights to the private sector for targeted company expansion.

Additional government initiatives, like the digitization of land and healthcare records by the Punjab Information Technology Board (PITB), illustrate the ability of big data to drive administrative effectiveness. The Safe Cities Project in Lahore utilizes surveillance cameras and data analysis to track crime and traffic, illustrating how technology can enhance public safety.

Global Frameworks in Data sharing & Data governance

Countries like Malaysia, Kenya, and Brazil have implemented data governance strategies that balance technological advancements with data privacy and regulation. Pakistan can learn from these experiences while adapting them to its socio-economic and legal landscape. Following are the several frameworks from developing & developed countries:[iv]

1. **Ghana Open Data Initiative (GODI)**

   Ghana Open Data Initiative (GODI) was started in January 2012 by the National Information Technology Agency (NITA) in partnership with the Web Foundation (WF), to make Government of Ghana data available to the public for re-use. It Promotes transparency, enables civic tech and policy.[v]

2. **DHIS2 (Africa)**

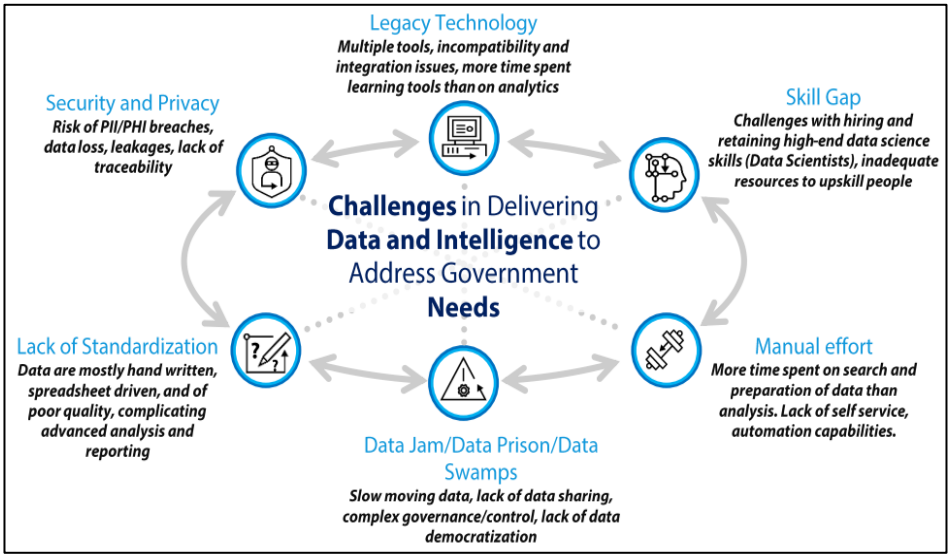
DHIS2 (District Health Information Software 2) is a widely used open-source, web-based platform for data collection, management, and analysis in health systems. It's particularly prevalent in Africa, with over 90% of African countries using it as a national-scale health information system. It strengthens health data, widely adopted for planning.[vi]

3. **NDSAP (India)**

The NDSAP (National Data Sharing and Accessibility Policy) in India is a government initiative that promotes open data access and sharing. It aims to make non-sensitive data, generated using public funds, available for lawful commercial and non-commercial use. It is an open government data, improved accessibility and innovation.[vii]

4. **U.S. & U.K Open Government Data Portals**

These government's official open data portals provide access to public sector datasets, aiming to improve government transparency, accountability, and citizen engagement. These portals provide access to datasets from major organizations.[viii]



Challenges in Adopting big data Practices in Pakistan

- Pakistan lacks the infrastructures that are needed to accommodate big data, such as data centers, cloud services, and high-speed internet, which are important for storing and processing data at a large scale.
- Pakistan is faced with erratic and incomplete data gathering, which hinders the use of big data. Most government departments and private sectors are not digitally mature to capture, hold, and update correct data sets.
- There is a dearth of data science, analytics, and big data technology professionals. Higher education institutions are just starting to introduce specialized courses, and there is not much availability of advanced training.
- Pakistan has not yet comprehensively developed its data protection legislation. The absence of proper guidelines for data privacy and usage can result in unwillingness by the business and organizational sectors to embrace complete big data strategies.
- Deploying solutions for big data entails major expenditure on infrastructure, software, and human resources. Most organizations, especially government agencies, are not willing to spend the necessary resources because of cost constraints.
- Data in Pakistan is often fragmented across different government departments and private entities. Lack of coordination and data sharing practices limit the ability to create unified, actionable insights.
- Big data systems are vulnerable to cyber-attacks. Pakistan's current cybersecurity infrastructure is still developing, making it challenging to protect large-scale data from breaches and other cyber threats.[ix]
-

Strategic Appraisal for strengthening Big Data Governance in Pakistan

Pakistan doesn't yet possess a data governance framework, national data strategy, national data ecosystem, and innovative data

technologies. In this regard, the following strategy is recommended: -

- **Establishing National Data Management Coordination Office**

  To begin with, Pakistan needs a national level central office for bringing in all involved stakeholders both in public and private sectors to join hands in implementing Data driven practices using Big Data technologies. Each ministry can implement Big Data services; the proposed office will provide general guidelines for all ministries to implement data related measures so that dream of data driven policy making can be achieved.[x]

Implementation Measures

- Ministry of Planning Development & Special Initiatives (MOPDSI) must spearhead the Data management coordination efforts because in the end this ministry is responsible for future planning of the country and it will be in need to approve development projects based on data analytics.
- To enhance transparency and avoid any political biases, proposed office must act as independent national body.
- Appropriate amendments to be introduced in constitution for its establishment.
- Proposed office must be comprised on representative of all stakeholders including private sector members, technology and legal experts, and academics.

- **Establishing Data Regulatory Governance Framework**

  First assignment for proposed coordination office must be to liaison with Ministry of law to introduce legal framework for data management by establishing data regulatory bodies at national and provincial levels. To meet set objectives this practice will require a major revision of existing disconnected data related SOPs among and within various departments.[xi]

Implementation Measures

- Pakistan should launch a Data Regulatory Governance Framework at national level with relevant constituent bodies like; Pakistan Telecommunication Authority (PTA), National Centre for Cyber Security (NCCS), Punjab Information Technology Board (PITB), Pakistan Space & Upper Atmosphere Research Commission (SUPARCO), etc. to combine and preserve structured data for futuristic automation.
- Ministry of Law along with MOPDSI to review existing legal provisions related to data handling.
- New provisions for data currency to be introduced as legal framework for National Data Management system.
- All departments to adopt data related SOPs based on this new legal framework to ensure uniform implementation of national data management framework across all the departments.
- The following table shows the regulations in different regions over the period of time with the key features of the laws.

Regulation	Region	Year	Key Features
GDPR (General Data Protection Regulation)	EU	2018	Enhanced privacy rights, global adoption, high penalties
CCPA (California Consumer Privacy Act)	California, USA	2020	Consumer rights, precedent for US privacy laws
DPA (Data Protection Act)	Kenya	2019	Improved awareness, implementation challenges
LGPD (Lei Geral de Proteção de Dados)	Brazil	2020	Similar to GDPR, strong consent requirements
POPI (Protection of Personal Information Act)	South Africa	2021	Comprehensive rights, accountable processing
PIPL (Personal Information Protection Law)	China	2021	Strict data localization, extraterritorial reach
PDPB (Personal Data Protection Bill)	Pakistan	2023	First comprehensive law, aligns with global standards

- **Data Ownership at National Level**

  Pakistan needs to take ownership of indigenous data. Global tech giants for capital

surveillance regulate the econometrics of Big Data. Both public and private sectors in Pakistan have large portfolios but lack sufficient coordination of Big Data and privacy issues in order to protect the data currency, Pakistan must develop a Centralized National Data Management and Surveillance system to enhance the efficiency and ownership of Big Data. This is the final strategic goal of the entire Big Data strategy and is critical for national security as well.[xii]

Implementation Measures

- Federal government must initiate a national data program with the strategic goal of keeping Pakistan's data on Pakistani data warehouses.
- To prevent legal issues about data privacy and protection, compliance with data protection legislation and best practices for data security is essential. Rules, regulations, and policies for handling big data at the national level under a central pool need to be implemented.
- Like the National Incubation Center, there is a need to establish a 'Big Data Industry' under a public-private partnership to cope with future trends of emerging technologies.
- Initially, Pakistan must focus on data domicile and data privacy requirements. Pakistan is not a small market by any measure. Government must negotiate with Tech giants to open offices in Pakistan.
- HEC through NCBC, Planning and IT ministries along must chart out a long term strategic vision to keep Pakistan's data local using 'Made in Pakistan' technologies.
- Based on these guidelines, ministry of planning with collaboration of HEC must identify HR requirements in various disciplines of related hardware and software technologies.
- Cabinet must declare it a strategic program with ensured funding.

➢ **Mechanism to Establish Framework for Big Data Governance**

- Comprehensive mechanisms to securely share and safeguard the integrity, confidentiality, and availability of massive organizational datasets with robust encryption protocols must be deployed to secure sensitive information, ensuring that only authorized entities have access.
- Stringent privacy regulations to govern the collection, storage, and utilization of big data, thereby establishing clear guidelines for responsible data management entities and advanced threat detection and prevention systems.
- There must be annual appraisal of required HR in various aspects of proposed National Data Management System.
- A performance approach based on clearly defined Key Performance Indicators (KPIs) should be adopted to evaluate the outcome of various stakeholders.
- Performance must be evaluated in short, mid and long term prospective after defining achievable milestones for each term.
- MOPDSI must hold responsible the stakeholders who miss their KPI targets.

- Evaluation process, once initiated, must continue without interruption.

Big Data Ownership and Future Course

In Nutshell, Pakistan is among top populated developing nations in the world who are facing unique challenges ranging from growing population, shrinking resources, climate change, and worsening security situation. In such circumstances, it is only natural to consider most optimal resource management policy. World's most successful economies are following data driven policies for optimal resource management. Pakistan unfortunately failed to adopt this model because of its comprehension challenge about the value of data in modern e-governance model, resource management, service delivery mechanisms, transparency, automation, economy and security.

Six government ministries of Pakistan including Ministry of Information Technology (MOIT), Ministry of Planning, Development and Reforms, National Database Registration Authority (NADRA), Ministry of Finance, Revenue and Economic Affairs, PTA and Ministry of Interior proposed a draft for the activation of big data industry in August 2016. Unfortunately, after many long years, since its inception there has been no visible action by government to active Big Data incubation across Pakistan except an initiative by HEC in 2018 when Higher Education Commission of Pakistan (HEC) in collaboration with 11 universities took Pakistan's first R&D initiative in Big Data and Cloud Computing and established National Centre for Big Data and Cloud Computing (NCBC) with an aim to develop novel Big Data solutions for various Pakistan specific requirements including taking ownership of Pakistan's public data from big tech giants Ministry of Planning Development & Special Initiatives (MOPDSI) is the largest stakeholder in data driven policy making in the country as it approves national development projects. Almost every government face criticism of having political

biases on approval of certain projects as existing method of project approval is not strictly based on data analytics and insights. The Planning Commission of Pakistan undertakes research studies and state policy development initiatives for the growth of national economy and the expansion of the public and state infrastructure of the country in tandem with the Ministry of Finance.

Similarly, MoIT is another big stakeholder in Pakistan related to data and its services along with national institutions like NADRA and FIA. Police, Power Distribution Companies, Surveillance systems, Hospitals etc. are all generating different kind of data on daily basis. This data driven approach requires embracing the latest data sciences innovations like Big Data, cloud computing and AI analytics so that data generated by national database management systems (NADRA, FIA), security apparatus (Safe City Cameras), social media and telecommunication networks can be collected, cleaned, categorized in a single national data management framework.

The ownership of big data is another critical aspect of global big data market. Whoever has the ownership of data holds significant economic and security leverage. In order to have the ownership of Pakistan's public data, it is critical to adopt a whole of nation approach. No single ministry will be able to implement it. This must be pursued as strategic projects as important as country's non-conventional deterrence projects. To begin with, NCBC's scope must be expanded and it must be declared as national level project instead of currently HEC's initiative. Recently, Pakistan's "One Nation One Vision" for 2025 advocates for modernization and institutional reform of the creation of a competitive knowledge economy and the public sector.[xiii]

The ability to store, manage, and utilise vast, unstructured, and heterogeneous data has become a pivotal aspect of modern digital capital, surpassing industrial innovations. Big data governance has evolved rapidly during the last 1 to 1.5 decade globally and its market is evolving at exponential rate which is in fact so high that many developed countries are struggling to keep pace with it. Big Data offers enormous opportunities in modern age ranging exploiting data currency. Major first world nations are harvesting benefits of Big Data while many developing nations are embracing it to overcome their governance challenges. A novel strategy has been proposed in this paper to overcome Pakistan's various governance and economic challenges stemming from lack of key data insights and their proper utilization which is only possible by embracing Big Data technologies.

Conclusively, Pakistan has not been fully equipped to capitalise on this innovation despite realizing its potential due to which Pakistan also lacks to adopt data driven policy mechanism in order to strategically manage resources, identify problem clusters, and enhance national productivity. Consequently, Pakistan's big data governance landscape is hindered by outdated methods, leading to poor assessment of socio-economic issues, inefficient resource allocation and ineffective policy implementation. While e-governance systems in various departments like; NADRA, FIA, and other utility based services have improved citizen service delivery, but adequately contribute at the policy level whereas the generation of big data is not utilised for future planning or innovation. The lack of centralised big data management strategy across public and private sectors is also encompassed with insufficient data compilation and concerns related to privacy concerns, further pose situation in the realm of cyber security and international surveillance from the big data giants. Ultimately, Pakistan reliance on global tech giants for data sharing results in economic and security concerns due to their unavoidable control, as the tech based

corporations exert in Pakistan's digital realm. By embracing this perspective and adopting forward-thinking strategies, Pakistan can effectively navigate the data revolution, unlocking its full potential for societal advancement and economic prosperity.

Reference

[i] Nasir, Jamal Abdul. "Big Data Prospects and Challenges for Pakistan." *The Nation*, https://www.nation.com.pk /15-Sep-2021/big-data-prospects-and-challenges-for-Pakistan, Accessed on 12 September 2024.

[ii] International Monetary Fund, https://www.imf.org/external/pubs/ft/fandd/2007/09/view.html Accessed on 14 September 2024.

[iii]Toward Big Data and analytics governance: Redefining Structural Governance Mechanisms https://www.researchgate.net/publication/348675261 Toward big data and analytics governance redefining structural governance mechanisms, Accessed on 20 September 2024.

[iv] https://www.pewresearch.org/short-reads/2018/06/04/how americans-have-viewed-government-surveillance-and-privacy-since-snowden-leaks/ Accessed on 22 October 2024.

[v]https://www.opengovpartnership.org/members/ghana/commitments/GH0019/ Accessed on 26 October 2024.

[vi] https://dhis2.org/ Accessed on 26 October 2024.

[vii] https://www.india.gov.in/national-data-sharing-and-accessibility-policy Accessed on 28 October 2024.

[viii] https://data.gov/open-gov/

[ix] https://zahid.quarto.pub/posts/data-government, Harnessing Data Analytics for Effective Governance in Pakistan: A Support, Not a Substitute. Accessed on 30 October 2024.

[x] Pas (2019) Big Data Challenges and opportunities, Pakistan Advertisers Society. Available at: https://pas.org.pk/6375-2/ Accessed on 02 November 2024.

[xi] Digital Regulation Platform. Available at: https://digitalregulation.org/navigating-data-governance-a-guiding-tool-for-regulators, Accessed on 15 December 2024.

[xii] UNESCAP, Available at: https://www.unescap.org/sites/default/d8files/event-documents/Malaysia-data-governance-profile.pdf, Accessed on 05 January 2025

[xiii] https://pu.edu.pk/downloads/Pakistan-Vision-2025.pdf, Pakistan Vision, One vision-One Nation, 2025, Accessed on 08 February, 2025.