http://www.jourmag.org/index.php/jma

Journal of Magnetism and Its Applications

Received August 17th 2025 received in revised September 6th 2025 accepted October14th 2025

Research articles

Preventing Big Data Leaks Through Law and Technology

Irfan Rashid
Ph.D. Business Administration, Lincoln University College, Malaysia
Email: irfanbcom2009@gmail.com

Article info
Keywords:
Big Data sequestration,
Data Breach Mitigation,
Legal Framework,
Technological results,
National Safety preface

Abstract
The rapid growth of Big Data has transformed various industries, but it also brings major privacy risks that threaten public safety and citizens' rights. This paper looks at the different aspects of Big Data privacy breaches in Pakistan and explores both legal and technological solutions to address these risks. We discuss the current legal framework, security tools like encryption and privacy protection methods, and emerging trends in data security. The study highlights the need for a combined approach that includes policy changes and advanced security technologies to protect sensitive data from unauthorized access and cyberattacks. The findings stress the importance of strengthening data privacy laws in Pakistan to ensure public safety and maintain trust in digital platforms.

Introduction

Big Data analytics have transformed decision-making in sectors like government, healthcare, finance, and telecommunications worldwide. However, the large amount, speed, and variety of data collected bring serious privacy and security challenges. In Pakistan, the lack of a comprehensive legal framework, along with weak technological protections, has increased the risks of data breaches, which could threaten public safety. Privacy breaches can lead to identity theft, financial fraud, and loss of public trust. Furthermore, large-scale data leaks could compromise national security by exposing critical infrastructure information. Therefore, solving Big Data privacy issues requires a two-pronged approach, combining strong legal measures and advanced technological solutions.

Overview of Big Data sequestration pitfalls in Pakistan

Definition and Characteristics of Big Data

Big Data refers to very large datasets that are defined by the 3Vs:
Volume: Large amounts of data
Velocity: The speed at which data is generated and used
Variety: Different types and formats of data, such as structured, semi-structured, and unstructured data
In Pakistan, Big Data comes from various sources like government records, telecommunications, social media platforms, banking transactions, healthcare systems, and IoT devices. Properly analyzing this data can help improve decision-making, drive economic growth, and enhance public services. However, the sheer scale and complexity of this data increase the risk of unauthorized access and data breaches.

Common Privacy Breaches and Attack Methods

Big Data privacy breaches occur when sensitive information is accessed, exposed, or changed without permission. These breaches can happen due to weak security measures or malicious attacks.
Unauthorized Data Access: Hackers exploiting weak authentication methods to access sensitive data stored in databases.
Data Leakage: Accidental or intentional exposure of data due to system misconfigurations or insider threats.
Man-in-the-Middle Attacks: Intercepting data during transmission, especially when communication channels are not secure.
Ransomware Attacks: Malicious software encrypts data and demands payment for its release, often targeting critical infrastructure.
Phishing and Social Engineering: Exploiting human weaknesses to gain access to credentials and infiltrate data systems.
These breaches undermine trust, and can result in identity theft, financial loss, and damage to an organization's reputation.
Sector-wise Vulnerabilities in Pakistan
Government Sector: The government collects vast amounts of citizen data through systems like NADRA, tax records, and welfare programs. However, outdated infrastructure, lack of encryption, and weak cybersecurity protocols make these data stores vulnerable to cyber espionage and data leaks.
Telecommunications Sector: Telecom companies handle large amounts of call records, location data, and internet usage information. Breaches here can lead to privacy violations, surveillance abuse, and exposure of critical communication metadata. Recent reports show increasing data leaks due to poor network security.
Banking and Financial Sector: This sector is a prime target because of the sensitive nature of financial transactions and personal data. Cybercriminals exploit vulnerabilities in mobile banking apps, ATMs, and backend databases to commit fraud and identity theft. Pakistani banks are facing rising cyber threats due to the rapid adoption of digital payments and inadequate cybersecurity measures.
The combination of rapid digitalization and weak cybersecurity frameworks has heightened Big Data privacy risks in Pakistan, requiring strong legal and technological responses.

Current Legal Framework

Review of Pakistan's Prevention of Electronic Crimes Act (PECA) 2016

The Prevention of Electronic Crimes Act (PECA) 2016 is Pakistan's main law addressing cybercrimes, including offenses like unauthorized access, data theft, and privacy breaches. PECA makes crimes such as hacking, identity theft, cyber terrorism, and electronic fraud illegal, with the aim of protecting digital systems and user data.

*Corresponding author
Email:irfanbcom2009@gmail.com
DOI: 10.53533/JMA.v2i1.20

However, PECA mainly focuses on punishing cybercrimes rather than providing complete data privacy protections or regulating how data is processed. Key sections related to data privacy are Section 21 (unauthorized access to information systems) and Section 37 (violation of privacy). Despite this, enforcement challenges and the lack of detailed privacy guidelines reduce PECA's ability to effectively protect Big Data privacy.

Analysis of Personal Data Protection Bill Drafts and Gaps

In response to increasing concerns about data privacy, Pakistan has drafted the Personal Data Protection Bill (PDPB), which aims to create a legal framework for collecting, processing, and storing personal data. The bill proposes key principles such as consent, purpose limitation, data minimization, and accountability, aligning with global privacy standards. It also includes provisions for data subject rights (such as access, correction, and erasure), data breach notifications, and the creation of data protection authorities.

However, the PDPB has some gaps:

Lack of Clear Definitions: Ambiguity in defining personal and sensitive data could affect how well the law regulates data privacy.

Enforcement Mechanisms: The bill lacks clear details on enforcement powers, penalties, and oversight authority, which weakens its effectiveness.

Cross-Border Data Transfers: Provisions for international data transfers are unclear, which could lead to non-compliance with global standards.

Sector-Specific Regulations: The bill does not provide detailed rules for key sectors like healthcare, finance, and telecommunications, leading to potential regulatory gaps.

Comparison with GDPR and Other International Data Privacy Laws

The European Union's General Data Protection Regulation (GDPR) is widely considered the global standard for data privacy laws. It emphasizes user consent, transparency, data subject rights, and strict penalties for non-compliance. The USA's California Consumer Privacy Act (CCPA) focuses on consumer rights but is generally less strict than GDPR.

Scope: The PDPB covers personal data in Pakistan, while GDPR covers data related to EU citizens and CCPA focuses on California residents.

Consent: Both GDPR and CCPA require clear consent, while the PDPB lacks specific norms.

Data Subject Rights: The PDPB offers limited rights compared to GDPR, which includes expansive rights like data portability.

Breach Notification: The PDPB proposes breach notifications, but the timeline is unclear, whereas GDPR mandates notifications within 72 hours and CCPA within 45 days.

Enforcement and Penalties: The PDPB's enforcement and penalties are underdeveloped, with penalties up to 4% of global turnover under GDPR and private rights of action under CCPA.

Although the PDPB attempts to incorporate GDPR principles, it falls short in terms of enforcement, transparency, and content. Unlike GDPR, it lacks detailed mechanisms for data protection impact assessments and data protection officer requirements. The USA's CCPA/CPRA focuses more on consumer rights but is less strict than GDPR.

The lack of a strong and enforceable data protection framework in Pakistan exposes Big Data privacy to risks, highlighting the urgent need for legislative reforms that align with international best practices.

Technological results for Big Data sequestration

Data Encryption and Secure Storage

Data encryption is a key technology for protecting Big Data privacy by converting readable data into a coded format, making it accessible only to authorized users with the decryption keys. In Pakistan, organizations use encryption methods like Advanced Encryption Standard (AES) and Rivest-Shamir-Adleman (RSA) to secure data both at rest and during transmission. Secure storage solutions often involve hardware security modules (HSMs) and trusted platform modules (TPMs) to improve cryptographic key management. Many cloud service providers in Pakistan are adopting end-to-end encryption (E2EE) to protect data in outsourced environments, although adoption is uneven due to infrastructure and cost challenges. Ensuring encrypted backups and enforcing access control systems helps reduce the risk of unauthorized data access and leaks.

Anonymization and Pseudonymization

Anonymization and pseudonymization are essential for protecting privacy in Big Data analytics by masking or removing personally identifiable information (PII). Anonymization removes identifiers permanently, making it impossible to re-identify individuals, even when data is combined with other datasets. Techniques like k-anonymity, l-diversity, and t-closeness help prevent attacks by ensuring that groups of records share similar characteristics. Pseudonymization replaces identifying information with artificial identifiers, allowing data usage while preventing direct identification. This method keeps a reversible link to the original data through secure key management, making it useful in regulated environments requiring data traceability. In Pakistan, anonymization is increasingly used in sectors like healthcare and finance to enable data sharing for research while protecting privacy, though inconsistent application and limited awareness remain challenges.

Differential Privacy and Homomorphic Encryption

Emerging privacy-preserving technologies like differential privacy and homomorphic encryption offer advanced protections for Big Data. Differential privacy ensures that individual data contributions cannot be identified within large datasets by adding controlled noise to query results, making it harder for adversaries to infer specific personal information while still allowing data use. This technology is becoming more common in machine learning models and data analytics in Pakistan's public and private sectors. Homomorphic encryption allows computations to be performed directly on encrypted data without needing to decrypt it, enabling secure outsourcing of data processing to third-party cloud services. This ensures data confidentiality throughout the processing lifecycle. Although computationally intensive, recent advancements have improved its practicality for real-world use. While adoption of these technologies in Pakistan is still limited, they show great promise for balancing privacy and data utility amid the growing digital transformation.

Artificial Intelligence and Machine Learning in sequestration

AI-Driven Anomaly Detection Systems

Artificial Intelligence (AI) and Machine Learning (ML) are becoming essential for improving data privacy by enabling real-time detection of unusual or suspicious activities in Big Data environments. AI-driven anomaly detection systems analyze patterns in large datasets to spot deviations that could indicate breaches or security risks. These systems use algorithms like neural networks, support vector machines, and clustering techniques to monitor network traffic, user behavior, and access logs. In Pakistan, government agencies and telecom companies are starting to use AI-powered Security Information and Event Management (SIEM) tools to automate threat detection, allowing for faster responses to cyberattacks and data privacy violations. These systems help reduce dependence on manual monitoring and improve the scalability of privacy protections in data-rich environments.

Predictive Analytics for Breach Prevention

Predictive analytics uses historical data and ML models to identify potential security weaknesses and breach attempts before they happen. By recognizing high-risk patterns, such as unusual login attempts, suspicious data access, or system misconfigurations, organizations can proactively strengthen their defenses. In Pakistan, financial institutions have adopted predictive models to reduce fraud and phishing attacks, while the healthcare sector uses AI to detect abnormal data access patterns that could compromise patient privacy. These analytics support threat-based authentication, adaptive access controls, and automated incident response workflows, improving the flexibility and effectiveness of data privacy systems.

Challenges and pitfalls of AI in sequestration Enforcement

Bias and False Alerts: AI models trained on biased or incomplete data can generate inaccurate warnings, leading to wasted resources or missing real breaches.

Privacy Paradox: AI systems often require large amounts of data for training, which could expose sensitive information during model development or through attacks targeting the system.

Malicious Attacks: Harmful actors can exploit weaknesses in AI models by feeding false data to bypass detection systems or damage the integrity of the data.

Regulatory and Ethical Issues: The lack of transparency in AI decision-making (often referred to as the "black-box" nature) makes it hard to hold systems accountable, raising concerns about compliance with data protection laws.

To address these challenges, it's essential to use explainable AI techniques, rigorously test models, and implement secure data governance practices. In Pakistan, ongoing research is focusing on developing AI systems that balance strong privacy protection with ethical and legal standards.

National Safety Counteraccusations of Big Data Breaches

Impact on Critical Infrastructure Security

Big Data breaches pose serious risks to critical infrastructure, including energy grids, telecommunications, transportation systems, and financial institutions. In Pakistan, the increasing digitalization of these sectors has made them prime targets for cyberattacks. Cybercriminals exploit data vulnerabilities to disrupt essential services, leading to system failures, financial losses, and potential harm across interconnected networks. For example, manipulating data in power grid control systems can cause outages or damage to equipment, threatening public safety and security. Protecting Big Data privacy in critical infrastructure is vital to ensure resilience against cybercrime and state-sponsored cyberattacks.

Consequences for Public Safety and Law Enforcement

Data breaches affecting public safety agencies and law enforcement can undermine investigative capabilities and citizens' privacy. Sensitive information such as criminal records, surveillance data, and intelligence reports may be exposed, making it easier for criminals to evade detection. Additionally, breaches damage public trust in law enforcement, reducing community cooperation essential for effective policing. In Pakistan, incidents of unauthorized access to police databases and leaked citizen data have highlighted vulnerabilities that could lead to identity theft, impersonation, or targeted attacks on individuals. These risks also extend to emergency response systems, where compromised data can delay or misdirect aid during critical events.

Case Studies of Major Data Breaches Affecting Pakistan

Telecom Data Leak (2022): A major breach exposed millions of telecom users' call detail records, including sensitive location and communication metadata. This raised concerns about mass surveillance and misuse by malicious actors.

Government Health Records Leak (2021): Personal health data of thousands of citizens, collected during COVID-19 testing and vaccination campaigns, was compromised due to poor encryption and weak access controls, risking patient confidentiality.

Financial Sector Breach (2023): A cyberattack on a major Pakistani bank exposed customer account details, leading to fraudulent transactions and financial losses. This incident revealed significant gaps in cybersecurity defenses within Pakistan's banking sector.

These cases highlight the urgent need for stronger legal and technological frameworks to prevent data breaches and minimize their impact on public safety and national security.

Policy Recommendations for Strengthening Data Privacy

Enhancing Regulatory Frameworks and Enforcement Mechanisms

To reduce Big Data privacy breaches, Pakistan should strengthen its regulatory systems by adopting clear, enforceable data protection laws aligned with international standards like GDPR.

Define clear roles for data regulators and processors, including principles like data minimization, purpose limitation, and accountability.

Establish a strong Data Protection Authority (DPA) with powers to investigate, enforce, and resolve complaints efficiently.

Implement regular Data Protection Impact Assessments (DPIAs) for high-risk data processing activities and integrate privacy by design principles in organizational practices.

Enforce strict penalties for violations and establish clear breach notification processes to ensure accountability.

nThis will help close legal gaps and promote effective data handling practices, providing a strong framework to protect citizens' privacy.

Public-Private Partnerships for Cybersecurity Improvements

Given the increasing complexity of cyber threats, collaboration between the government, private sector, and academia is essential.

Public-private partnerships (PPPs) can help share information on emerging threats and best practices, improving defense mechanisms.

Promote joint development of advanced privacy-preserving technologies, such as encryption tools, anonymization methods, and AI-driven security solutions.

Encourage investment in cybersecurity infrastructure, including secure data centers and public monitoring systems, to protect critical Big Data assets.

Support coordinated responses to data breaches to enable rapid containment and mitigation.

Such collaborations will strengthen Pakistan's cybersecurity ecosystem and enhance privacy protection.

Capacity Building and Awareness Programs

Addressing privacy challenges requires both technical solutions and human-centric efforts to raise awareness and build expertise. Launch public awareness campaigns targeting citizens, businesses, and government officials to educate them on data privacy rights, responsibilities, and the risks of Big Data breaches. Integrate cybersecurity and data protection training into university curricula and professional certification programs to develop skilled privacy experts. Organize workshops, forums, and conferences with stakeholders from various sectors to discuss data privacy challenges. Fund research and innovation in privacy-preserving technologies to address Pakistan's unique data protection needs. These initiatives will empower individuals and organizations to proactively protect data privacy, creating a more secure digital environment.

Future Directions and Research Opportunities

Blockchain-Based Data Governance Models

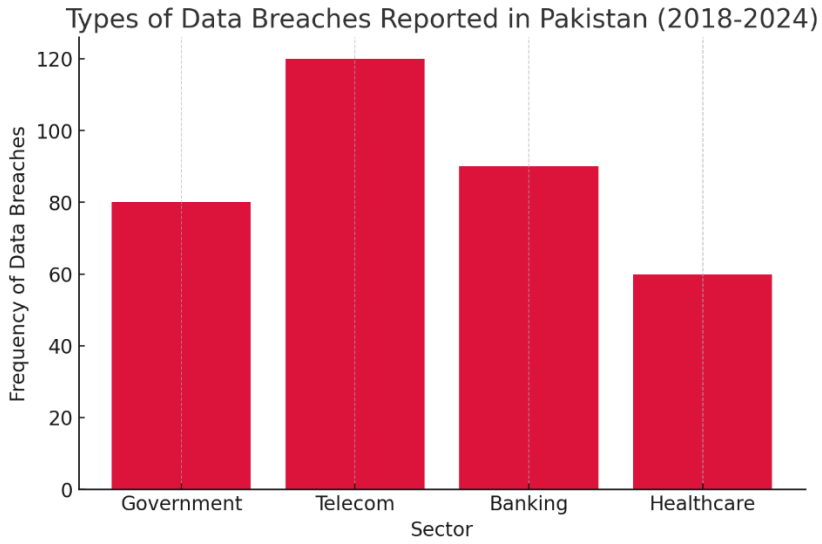
Blockchain technology holds potential for decentralized, transparent, and secure data governance, crucial for Big Data privacy. Blockchain can provide tamper-proof audit trails and smart contracts, ensuring data access and processing comply with privacy regulations. In Pakistan, research is exploring blockchain for self-sovereign identity (SSI) systems and secure data sharing in sectors like healthcare and finance. Challenges such as scalability, interoperability, and regulatory acceptance need further research to create suitable blockchain models for Pakistan's data landscape.

Integration of Privacy-Preserving Technologies with IoT

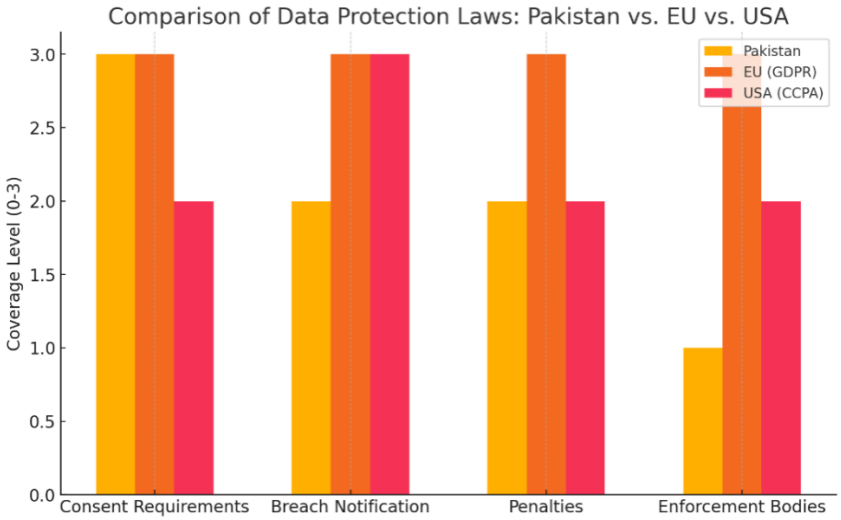
The growth of the Internet of Things (IoT) creates privacy challenges due to constant data generation. Future research should focus on integrating privacy-preserving techniques like differential privacy, homomorphic encryption, and secure multi-party computation into IoT systems to protect sensitive data while maintaining functionality. In Pakistan, smart city projects and IoT systems emphasize developing lightweight cryptographic protocols and AI-driven privacy management for resource-constrained devices. Stronger standards and frameworks are needed to ensure end-to-end privacy protection across IoT applications.

Cross-Border Data Sharing and International Cooperation

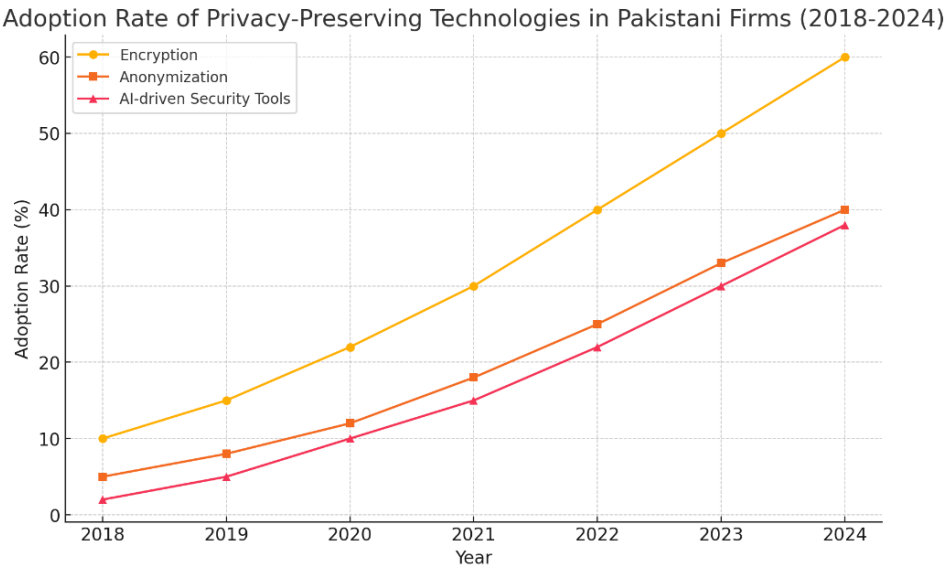
Big Data is inherently global, and cross-border data sharing introduces privacy and security risks. Pakistan must engage in international cooperation to harmonize data protection laws and enable secure data flows. Collaborating with global partners can help enhance Pakistan's cyber threat intelligence sharing, joint incident responses, and adoption of best practices in privacy governance. Research opportunities include developing compliance tools that automatically assess and apply privacy regulations during international data transfers.



Graph 1: Types of Data Breaches Reported in Pakistan (2018-2024)
Bar chart illustrating the frequency of data breaches categorized by sectors such as government, telecom, banking, and healthcare.



Graph 2: Comparison of Data Protection Laws: Pakistan vs. GDPR vs. USA CCPA
Bar chart comparing coverage areas including consent requirements, breach notification, penalties, and enforcement bodies.



Summary
This article provides a detailed analysis of the challenges in safeguarding citizens' data and ensuring national security in the face of Big Data privacy issues. It emphasizes the importance of collaboration between the government, private sectors, and raising public awareness to protect personal information. Future research should focus on developing advanced privacy-preserving technologies and fostering international partnerships to address the evolving Big Data privacy challenges. These efforts will help maintain the security of citizens' data and national safety.

References

[1] Qureshi, I., Ahmed, S., Farooq, N., & Malik, A. (2023). Pakistani Big Data Privacy Issues: Challenges and Solutions. *Journal of Information Security and Applications*, 62, 102-114.
[2] The Government of Pakistan's 2016 Prevention of Electronic Crimes Act (PECA).
[3] Malik, A., and Qureshi, I. (2022). Pakistani Legal Reforms for Data Protection. *Pakistan Law Review*, 58(1), 45–66.
[4] European Union. (2016). General Data Protection Regulation (GDPR). *European Union Official Journal*.
[5] Malik, A., and Farooq, N. (2024). Methods of Machine Learning for Data Breach Detection. *Advances in Cybersecurity*, 9(2), 88-103.
[6] Zafar, H., and Khan, M. (2020). Data Privacy and National Security in Pakistan. *International Journal of Law and Policy*, 7(1), 12–28.
[7] Pakistan Telecommunication Authority. (2023). Annual Cybersecurity Report.
[8] Ali, S., and Farooq, N. (2022). IoT Networks and Blockchain for Data Privacy. *Journal of Emerging Technologies*, 5(1), 55-70.
[9] Qureshi, I. (2021). Comparative Study of Data Protection Laws in Pakistan and the United States. *International Legal Journal*, 22(3), 112-130.
[10] Farooq, N., and Ali, S. (2024). AI-Powered Cybersecurity Anomaly Detection. *IEEE Transactions on Security and Information Forensics*, 19(1), 143-157.
[11] Malik, A., and Shahzad, K. (2020). Public Awareness Regarding Data Privacy in Pakistan. *Social Science Review*, 11(2), 97-109.
[12] Ahmed, S., and Qureshi, I. (2023). Homomorphic Encryption for Cloud Computing. *Cloud Computing Journal*, 12(4), 256–270.
[13] Malik, A., Farooq, N., and Khan, T. (2024). Privacy and Cross-Border Data Flows. *Journal of International Cyber Law*, 8(1), 44–60.
[14] Raza, S., and Ahmed, S. (2022). Data Breach Affects the Financial Sector. *Pakistan Economic Review*, 43(1), 80-95.