



Dr. Gulzar Ahmad¹, Junaid Shaukat² & Muhammad Faraz Tahir Bhatti³

¹Assistant Professor, Riphah International University, Islamabad.

²Research Scholar, Masters of Computer Science, Pir Mehr Ali Shah Arid Agriculture University, Rawalpindi.

Email: junaidshaukat042@gmail.com

³Research Scholar, BS Information Technology, Minhaj University Lahore. Email: farazbhatti140@gmail.com

KEYWORDS	ABSTRACT
Pakistan; cybersecurity; cloud computing; artificial intelligence; Zero Trust Architecture; National Cyber Security Policy; critical information infrastructure; incident response	Pakistan's expanding use of digital public services, online banking, telecommunications, cloud platforms, and connected infrastructure creates opportunities for economic development while increasing exposure to cyber risk. Conventional perimeter-based security is increasingly insufficient where employees, contractors, applications, and data operate across distributed networks and third-party platforms. This conceptual paper examines how cloud-based artificial intelligence (AI) and Zero Trust Architecture (ZTA) could strengthen Pakistan's cybersecurity capabilities. It draws on a narrative review of Pakistan's National Cyber Security Policy 2021, the Computer Emergency Response Team (CERT) Rules 2023, the National AI Policy 2025, recent statements by the national CERT (PKCERT), State Bank of Pakistan (SBP) technology-risk guidance, and global evidence from industry and standards bodies, including the NIST Zero Trust Architecture, the CISA Zero Trust Maturity Model, and the IBM and Verizon breach studies. The paper explains how AI-enabled analytics can support anomaly detection, threat prioritisation, and incident response, while ZTA can enforce identity verification, least-privilege access, device assessment, and network segmentation. It documents the recent threat picture reported for Pakistan, identifies context-specific implementation concerns (skills shortages, uneven organisational maturity, the absence of a comprehensive data protection statute, legacy infrastructure, third-party cloud dependence, adversarial attacks on AI, and incident-reporting coordination), and proposes a five-stage strategic framework with measurable indicators for public institutions, financial services, telecommunications, critical infrastructure operators, and small and medium-sized enterprises (SMEs). The paper is conceptual and does not claim to measure the effectiveness of deployed systems in Pakistan. It concludes that AI and ZTA should be introduced through risk-based pilots, clear accountability, workforce development, independent testing, and alignment with Pakistan's evolving policy and regulatory environment.
ARTICLE HISTORY Date of Submission: 24-02-2026 Date of Acceptance: 28-03-2026 Date of Publication: 31-03-2026	 2026 Journal of Social Sciences Development
Corresponding Author	Junaid Shaukat
Email:	junaidshaukat042@gmail.com
DOI	https://doi.org/10.59075/jssd.v5i1.402
Volume/Issue	5(1), 2026, 392-412

Introduction

Pakistan's evolving digital security environment

Digital transformation is changing how Pakistani public institutions, banks, telecommunications providers, universities, businesses, and citizens access and exchange information. DataReportal estimated that Pakistan had about 117 million internet users at the end of 2025, equivalent to 45.6% of the population (DataReportal, 2025). Online services and cloud-hosted applications can improve availability and efficiency, but they also create dependencies on digital identities, network connectivity, software suppliers, data centres, and service providers. A compromise affecting one account or third-party service may therefore create risks beyond the organisation in which the incident begins.

Cybersecurity in this environment should be understood as a national resilience issue rather than solely a technical IT function. Government records, financial transactions, telecommunications, energy operations, health information, and other essential services depend on the confidentiality, integrity, and availability of digital systems. Ransomware, credential theft, phishing, data exposure, malicious software, misconfigured cloud services, and attacks on suppliers can disrupt services or undermine public confidence. The likelihood and consequences of specific threats vary by sector and organisation; risk assessments should therefore be based on evidence rather than assumptions about a single uniform national threat profile.

Pakistan has established policy and institutional foundations relevant to this challenge. The National Cyber Security Policy 2021 provides a national policy reference, while the Computer Emergency Response Team Rules 2023 describe a multi-tier CERT arrangement covering national, sectoral, and organisational responsibilities, including coordination for critical information infrastructure (Government of Pakistan, 2023; Ministry of Information Technology and Telecommunication [MoITT], 2021). In the financial sector, SBP guidance addresses technology risk governance, cybersecurity controls, and the management of risks associated with digital financial services and third-party arrangements (State Bank of Pakistan [SBP], 2025, n.d.). More recently, the National AI Policy 2025 has added a national position on AI-based cybersecurity solutions and secure AI ecosystems (MoITT, 2025), and PKCERT has announced a requirement for organisations to establish security operation capabilities (Business Recorder, 2026; Hanif, 2026). These instruments create a basis for operational improvement, but policy objectives must be translated into measurable controls, adequate staffing, tested response procedures, and regular assurance.

Problem statement

Many organisations still rely on a combination of perimeter firewalls, antivirus tools, periodic audits, and manual incident handling. These controls remain important, but they may not provide sufficient visibility across cloud environments, remote access, mobile endpoints, interconnected suppliers, and complex identity systems. Attackers who obtain valid credentials may be able to

move through an environment without triggering traditional perimeter alerts. The 2025 Verizon Data Breach Investigations Report, which analysed more than 22,000 incidents and 12,195 confirmed breaches worldwide, found credential abuse (22%) and vulnerability exploitation (20%) to be the leading initial access vectors, and reported that third-party involvement in breaches had doubled to 30% (Verizon, 2025). At the same time, security teams can face high alert volumes and limited capacity to investigate every event promptly.

AI and ZTA offer complementary capabilities: AI can help identify unusual patterns across large volumes of telemetry, while ZTA reduces implicit trust and limits the access available to any user, device, or service. However, neither technology is a complete solution. AI can produce false positives, inherit bias from data, or be manipulated; ZTA can be undermined by poor identity governance, excessive exceptions, or incomplete asset inventories. Pakistan therefore needs an implementation approach that considers technical feasibility, cost, governance, privacy, skills, and regulatory alignment together.

Aim, objectives, and contribution

The aim of this paper is to develop a Pakistan-oriented conceptual framework for combining cloud-based AI and Zero Trust Architecture to strengthen organisational and national cyber resilience. The objectives are:

- To explain the complementary roles of cloud-based AI and ZTA in modern cybersecurity.
- To summarise the recent cyber-threat evidence reported for Pakistan and interpret its limitations.
- To relate these technologies to Pakistan's policy and institutional context, including recent developments in 2025–2026.
- To identify technical, organisational, financial, legal, and ethical barriers to adoption.
- To propose a phased implementation framework, with indicators, and practical recommendations for key sectors.
- To outline research questions for future empirical evaluation.

The paper contributes in three ways. First, it consolidates scattered public reporting on Pakistan's threat environment and makes explicit why the available figures cannot be compared directly. Second, it links the most recent national directives (including the PKCERT security-operations requirement) to a staged AI-ZTA adoption path. Third, it translates international guidance into sector-specific priorities and testable indicators suited to Pakistan's institutional conditions.

Method and scope

This paper uses a conceptual, narrative-review approach. It synthesises the main technology themes in the initial source article (Kolawole, 2025) with Pakistan-specific official policy and regulatory material, including the National Cyber Security Policy 2021, the CERT Rules 2023,

the National AI Policy 2025, and SBP technology-risk publications. It also draws on the NIST Zero Trust Architecture publication, the NIST AI Risk Management Framework, the NIST Cybersecurity Framework 2.0, and the CISA Zero Trust Maturity Model to structure the discussion of access control, governance, and responsible AI (CISA, 2023; National Institute of Standards and Technology [NIST], 2023, 2024; Rose et al., 2020).

Evidence on the current threat environment was collected from publicly available sources in October 2026. These include statements by PKCERT and the National Cyber Crime Investigation Agency (NCCIA) as reported in the national press, telemetry published by a commercial security vendor (Kaspersky, as reported in Pakistani media), and international studies by IBM, Verizon, and ISC2. Three limitations of this evidence base should be noted. Much of the Pakistan-specific data reaches the paper through news reporting of conference presentations and press briefings rather than through peer-reviewed or primary statistical publications. Vendor telemetry reflects the vendor’s own customer base and detection definitions. Finally, official incident counts and vendor detection counts measure different things and differ by orders of magnitude; they are reported here side by side only to illustrate the range of available indicators, not to estimate a single national attack rate.

The paper does not report a survey, experiment, penetration test, or original incident dataset. Accordingly, it makes no statistical claims about the prevalence of attacks in Pakistan or quantified improvements from AI or ZTA in Pakistani organisations. The proposed framework is a reasoned set of recommendations that should be tested through organisational case studies, interviews, maturity assessments, and controlled pilots. Before journal submission, the author should verify the current legal status and latest versions of all policies and regulatory instruments, particularly those described as draft or under review, and conduct a systematic literature search with documented inclusion criteria.

The reported cyber-threat picture in Pakistan

Scale and composition of reported activity

Publicly reported figures on Pakistan’s cyber incidents come from several institutional vantage points, which helps to explain why they diverge. Table 1 summarises the main indicators identified for this paper together with their sources and an interpretive note.

Table 1 Reported cyber-threat and cybersecurity indicators for Pakistan

Indicator		Reported figure	Source type	Interpretive note
On-device detections, 2025	attack Jan-Sep	More than 5.3 million; 126,000 spyware; 42,000 ransomware	Vendor telemetry (Kaspersky), ProPakistani and Daily Times (2025)	Detections, not confirmed intrusions; reflects vendor user base

Incidents handled by national CERT	927 in 2024–2025 combined; 253 in 2026 by early July	PKCERT statements (Business Recorder, 2026; Hanif, 2026)	Reported or handled incidents; likely a lower bound
Cybercrime complaints	More than 73,000 complaints in 2024; 1,604 cases registered	FIA annual report, via Profit (2025)	Large gap between complaints and registered cases
Change in cybercrime in 2025	Increase of 35%, driven by WhatsApp hacking and financial fraud	NCCIA officials, via Profit (2025)	Official characterisation; final 2025 figure may differ
National cybersecurity standing	Tier 1 (role-modelling), ITU GCI 2024; previously 79th	MoITT/PTA statement (Pakistan Telecommunication Authority, 2024)	Measures national measures, not organisational security

Note. Figures are as reported in the cited sources and use different definitions; they should not be summed or compared directly.

Kaspersky reported that its products detected more than 5.3 million on-device attacks in Pakistan during the first three quarters of 2025, including about 126,000 spyware detections, 113,000 backdoors, 107,000 password stealers, and 42,000 ransomware incidents (ProPakistani, 2025). The same reporting stated that 27% of Pakistani users and 24% of corporate entities encountered malware through local sources such as removable media and hidden installers. A separate report of the same briefing noted more than 2.5 million web-based attacks and 166,000 banking-malware detections, and linked exposure to outdated systems and poor cyber hygiene (Daily Times, 2025). These are detection counts from one vendor’s telemetry, not confirmed intrusions.

By contrast, PKCERT’s Director General reported 927 cyberattacks in 2024 and 2025 combined and 253 so far in 2026, including a ransomware attack on an unnamed government entity that encrypted data and forced restoration from backups; he cautioned that these were the incidents successfully handled and that many organisations remain unaware of breaches (Business Recorder, 2026; Hanif, 2026). The much smaller number reflects incidents that were reported to, or handled by, the national CERT. Because incident reporting depends on organisational awareness and willingness to report, the official count is best interpreted as a lower bound on significant incidents rather than a measure of total attempted attacks.

Fraud, social engineering, and the human element

Complaint-based indicators point to a threat environment dominated by fraud and social engineering as well as by technical attacks. According to reporting based on the Federal Investigation Agency’s 2024 annual report, more than 73,000 cybercrime complaints were filed in Pakistan in 2024, but only 1,604 cases were formally registered, and NCCIA officials said that cybercrime had risen by 35% in 2025, driven by WhatsApp account takeover, online scams, and

financial fraud (Profit, 2025). Officials attributed this partly to low digital literacy and the ease with which victims could be persuaded to disclose one-time passwords. The same report cited Visa findings that every second Pakistani experienced online financial fraud in 2024.

These patterns matter for the AI-ZTA argument in two ways. First, many successful attacks begin with the compromise of a person or an identity rather than with exploitation of a network perimeter, which is the situation in which identity-centric controls such as multifactor authentication, conditional access, and session risk scoring have the greatest value. Second, consumer-facing fraud and organisational compromise are connected: stolen credentials, hijacked messaging accounts, and fraudulent calls can be used to reach employees, suppliers, and customers of organisations that are themselves reasonably well defended.

A policy–practice gap

Pakistan's formal cybersecurity standing has improved. The Ministry of IT and Telecommunication and the Pakistan Telecommunication Authority announced that Pakistan had been placed in Tier 1 ("role-modelling") of the International Telecommunication Union's Global Cybersecurity Index 2024, up from 79th position in the previous edition, citing the National Cyber Security Policy, the CERT Rules, the Critical Telecom Data and Infrastructure Security Regulations, the national telecom security operations centre, and regular audits of the telecom sector (Pakistan Telecommunication Authority, 2024). Yet officials quoted in the press describe a disconnect between this improved framework and everyday practice, attributing continuing incidents to limited public awareness and unsafe digital behaviour (Profit, 2025). The Global Cybersecurity Index measures legal, technical, organisational, capacity-building, and cooperation measures at national level; it does not measure the security of individual organisations. A high index tier is therefore compatible with uneven control maturity inside institutions, which is the gap that operational frameworks such as AI-supported monitoring and Zero Trust are intended to address.

Cloud-based AI for cybersecurity

Capabilities and potential applications

Cloud-based AI combines scalable computing and storage with machine-learning and analytical techniques that help security teams process data from network devices, endpoints, applications, identity providers, and cloud workloads. Models can learn patterns associated with ordinary activity and flag deviations for investigation. Examples include an unusual sign-in location, a sudden increase in failed logins, atypical data downloads, unexpected changes to privileged accounts, or communication between systems that rarely interact.

For Pakistani organisations, the most useful initial applications are likely to be those with clearly defined operational outcomes: prioritising security alerts, detecting account compromise, identifying suspicious email or endpoint behaviour, correlating logs from multiple systems, and supporting incident triage. These uses fit directly with the new PKCERT expectation that

organisations operate security operation centres supported by tools such as security information and event management (SIEM), endpoint detection and response (EDR), network detection and response (NDR), and extended detection and response (XDR) (Hanif, 2026). Automated playbooks may isolate a device, revoke a session, or block a known malicious indicator, but high-impact actions should have safeguards, audit trails, and human approval where appropriate. AI outputs should inform decisions rather than replace accountable security professionals.

What the international evidence suggests

The strongest large-sample evidence on the financial value of AI-supported security comes from the IBM Cost of a Data Breach Report 2025, based on research by the Ponemon Institute. It reported a global average breach cost of US\$4.44 million, the first decline in five years, and an average breach lifecycle (time to identify and contain) of 241 days. Organisations that used AI and automation extensively across their security operations saved an average of US\$1.9 million per breach and shortened the breach lifecycle by about 80 days relative to organisations that did not (IBM, 2025). Organisations that detected breaches internally rather than being notified by an attacker also saved about US\$900,000. These figures are associational, drawn from organisations that had already experienced breaches, and mainly from high-income economies; they cannot be transferred directly to Pakistani organisations, whose cost structures, data volumes, and regulatory exposure differ. They do, however, support the hypothesis that faster detection and containment is the primary mechanism through which AI-assisted security reduces harm.

The same study carries a caution that is directly relevant to Pakistan's National AI Policy ambitions. IBM reported that 13% of organisations had experienced a breach involving an AI model or application, and that almost all of these lacked proper AI access controls; it also found that "shadow AI" (unsanctioned use of AI tools) added about US\$670,000 to average breach costs where usage was high (IBM, 2025). AI adoption can therefore create new attack surface at the same time as it improves defence. Governance of AI tools used by staff, and of models embedded in security products, is a requirement rather than an optional extra.

Benefits and limitations

Cloud platforms can make advanced security analytics more accessible to organisations that cannot build large in-house computing environments. Shared threat intelligence and centralised monitoring may also improve consistency across dispersed sites. Nevertheless, cloud adoption introduces questions about data location, provider access, service availability, contractual responsibilities, vendor concentration, and the ability to export logs during an incident. Organisations should assess provider security, subcontractors, encryption, retention, access logging, recovery arrangements, and exit options before moving sensitive workloads.

AI performance depends on the quality, coverage, and representativeness of the available data. Poorly tuned models may overwhelm analysts with false alarms or fail to detect new attack patterns. Models trained on data from other regions may also encode assumptions about normal

behaviour that do not hold in Pakistani workplaces, for example in relation to shared devices, shared accounts, prepaid mobile connectivity, or working-hour patterns. Models can be deliberately manipulated: NIST's taxonomy of adversarial machine learning describes evasion, poisoning, privacy, and abuse attacks against predictive and generative AI systems, and notes that no single mitigation is complete (Vassilev et al., 2025). Opaque recommendations can also make incident decisions difficult to explain. Governance should therefore include documented model purpose, data minimisation, validation against realistic scenarios, drift monitoring, periodic reassessment, and a clear process for challenging or overriding automated decisions (NIST, 2023).

Zero Trust Architecture and its relevance to Pakistan

Core principles

Zero Trust Architecture does not treat a user or device as trustworthy merely because it is connected to an internal network. Access decisions should consider identity, device condition, requested resource, context, and risk. NIST's ZTA guidance describes a shift away from static network-perimeter assumptions toward explicit, continuously evaluated access decisions (Rose et al., 2020). The practical principles include verifying each access request, applying least privilege, using multifactor authentication, managing privileged accounts carefully, segmenting networks, recording access activity, and removing access when it is no longer justified. NIST has also published guidance on applying zero trust access control to cloud-native applications deployed across multiple locations, which is relevant to organisations using hybrid or multi-cloud arrangements (Chandramouli & Butcher, 2023).

The CISA Zero Trust Maturity Model organises implementation across five pillars (identity, devices, networks, applications and workloads, and data) supported by cross-cutting capabilities for visibility and analytics, automation and orchestration, and governance, and it describes progression through traditional, initial, advanced, and optimal stages (CISA, 2023). The maturity-stage logic is useful for Pakistan because it treats Zero Trust as a journey that can begin with modest controls rather than as a product to be purchased in a single step. Table 2 maps the pillars to practical starting points for Pakistani organisations and to the role AI analytics can play.

Table 2 Mapping of Zero Trust pillars to Pakistani implementation starting points

Pillar (CISA)	Practical starting point in Pakistan	Potential role of AI analytics
Identity	Identity inventory; MFA for privileged and remote access; removal of dormant accounts	Detect anomalous sign-ins and privilege changes; risk-score sessions
Devices	Asset register; endpoint protection; device-compliance checks for sensitive systems	Flag unmanaged or compromised devices; endpoint behaviour analytics
Networks	Segmentation of sensitive workloads; controlled remote and vendor access	Detect unusual east-west traffic and rare system-to-system communication

Applications and workloads	Secure cloud configuration; least-privilege application roles; patching of exposed services	Identify misconfiguration drift and abnormal workload behaviour
Data	Data classification; retention rules; protected backups; access logging	Detect atypical downloads and exfiltration patterns
Cross-cutting: visibility, automation, governance	Central logging; documented playbooks; executive ownership; exception handling	Correlate signals; trigger approved automated responses with audit trails

Note. Pillars follow the CISA Zero Trust Maturity Model (CISA, 2023); starting points and AI roles are the author’s proposals.

Relevance to Pakistani organisations

In Pakistan, ZTA could be especially relevant to organisations with multiple offices, remote staff, outsourced IT, cloud applications, legacy systems, or sensitive citizen and customer data. Public-sector departments can begin with identity inventories and multifactor authentication for privileged and remote access. Banks and payment institutions can strengthen identity and device controls around administrative access and high-risk operations. Universities and SMEs can use managed identity services and role-based access as an achievable starting point, provided that access rights are regularly reviewed.

The case for an identity-first emphasis is supported by the global breach data. The 2025 Verizon report found credential abuse to be the most common initial access vector and reported that the proportion of breaches involving third parties had doubled, with ransomware affecting small and medium-sized businesses disproportionately (Verizon, 2025). Organisations in Pakistan that depend on outsourced IT, software resellers, and managed service providers face a comparable exposure, and ZTA’s emphasis on time-limited, monitored, least-privilege access for external parties directly addresses it.

Integrating AI and ZTA

AI and ZTA address different but complementary parts of cyber risk. AI helps interpret signals and identify suspicious activity; ZTA enforces access decisions and limits what a compromised identity can reach. For example, an AI system might flag a user’s unusual access pattern. A policy engine could then require stronger authentication, restrict the session, or deny access to a sensitive resource while an analyst investigates. This approach is most effective when asset inventories, identity records, event logs, and response procedures are reliable.

Integration should be incremental. An organisation should not automate broad account suspensions or network isolation before it understands normal operations and has tested recovery procedures. Policies need exception handling, emergency access, separation of duties, and an auditable record of both automated and human decisions. A zero-trust programme should be

evaluated by reduced excessive access, improved visibility, and resilience, not by the number of tools purchased. The NIST Cybersecurity Framework 2.0, which adds a dedicated Govern function alongside Identify, Protect, Detect, Respond, and Recover, offers a convenient vocabulary for organising these activities and for communicating them to executives and regulators (NIST, 2024).

Pakistan's policy and sectoral context

National cyber governance and CERT coordination

Pakistan's National Cyber Security Policy 2021 is an important national reference for cybersecurity governance (MoITT, 2021). The Computer Emergency Response Team Rules 2023 set out a layered arrangement involving national, government, critical-infrastructure, sectoral, provincial, and organisational CERT functions. The Rules include proactive activities such as threat analysis, advisories, vulnerability management, security assessments, and coordination, as well as responsive activities such as incident handling, investigation, recovery, and lessons learned (Government of Pakistan, 2023).

Two developments in 2026 sharpen the operational expectations. First, PKCERT's Director General announced that all public and private organisations had been given six months to establish fully operational security operation centres for continuous threat monitoring, and that a federal-government CERT would be established under the national CERT (Business Recorder, 2026; Hanif, 2026). Because the announcement was made in early July 2026, the deadline falls around the start of 2027. Second, the Pakistan Information Security Framework 2026 (PISF 2026), described by PKCERT as containing compliance requirements and audit controls and as the country's first national cybersecurity standard, was reported to be under cabinet review and presented to parliament. Its final text and legal status should be verified before the framework proposed here is aligned to it. For many organisations, particularly smaller ones, the practical question is how to meet a monitoring requirement without building a full 24-hour in-house operation. Managed detection, shared sectoral centres, and cloud-based AI analytics are plausible answers, and they are the context in which the staged framework in Section 8 is offered.

The proposed AI-ZTA approach should complement these structures rather than create disconnected reporting channels. Organisations should maintain an incident-response plan that identifies who classifies incidents, who contacts the relevant CERT or regulator, what evidence must be preserved, how suppliers participate, and how recovery is approved. Shared threat indicators can be useful, but information-sharing arrangements should define classification, privacy protections, authorised recipients, and procedures for correcting inaccurate indicators.

National AI policy and cyber law

The federal cabinet approved Pakistan's first National AI Policy on 30 July 2025 (Dawn, 2025). The policy is organised around six pillars, including a secure AI ecosystem that introduces

regulatory sandboxes, cybersecurity protocols, and transparency frameworks, and an AI infrastructure pillar that envisages a national compute grid and cloud-based resources; it also sets a goal of training one million AI professionals by 2030 (Business Recorder, 2025; MoITT, 2025). The policy's contents on AI-based cybersecurity solutions and on trust and identity management are directly relevant to the framework proposed here. They also create an opening for pilot projects to be designed within the policy's sandbox and capacity-building mechanisms, subject to implementation details that were not yet public at the time of writing.

In the criminal-law domain, the 2025 amendments to the Prevention of Electronic Crimes Act created new bodies including the National Cyber Crime Investigation Agency, the Social Media Protection and Regulatory Authority, and a tribunal, according to legal commentary (DLA Piper, n.d.). These bodies concern enforcement and online-content regulation more than organisational security standards, but they shape the reporting and investigative environment in which security incidents are handled.

Financial services and digital payments

Financial services depend on identity systems, transaction platforms, customer-facing applications, third-party service providers, and continuous availability. SBP's published framework for payment institutions highlights technology governance and cyber-risk management, with implementation requirements proportionate to the size, nature, and complexity of the institution (SBP, 2025). SBP's broader supervisory materials also identify cybersecurity, technology risk, third-party risk, and operational resilience as important supervisory areas (SBP, n.d.).

AI-based transaction and access analytics may help detect suspicious behaviour, while ZTA can strengthen privileged access, administrative sessions, and access to critical systems. These controls should be integrated with fraud monitoring, business continuity, customer protection, and regulatory reporting. Given the scale of reported consumer fraud described in Section 3.2, the combination of device-binding, step-up authentication for high-risk transactions, and behavioural analytics is particularly relevant to digital banking and mobile wallets. Financial institutions should not assume that a cloud provider or AI vendor transfers their accountability for customer data and service resilience.

Public services, telecommunications, and critical infrastructure

Public-sector systems may hold sensitive personal records and support essential administrative functions. The June 2026 ransomware incident at a government entity illustrates both the exposure and the value of tested backups, since recovery from backups was reported as the means of restoring data (Hanif, 2026). Telecommunications, energy, transport, and other infrastructure sectors can also depend on operational technology (OT) whose availability and safety requirements differ from ordinary office IT. The CERT Rules explicitly recognise critical information infrastructure and include sectoral arrangements for areas such as telecom, banking,

finance, power, transport, and public services (Government of Pakistan, 2023). The telecom sector in particular has a national telecom CERT and a national telecom security operations centre, which could serve as a model for sectoral shared services (Pakistan Telecommunication Authority, 2024).

For operational technology, AI and ZTA controls must be introduced carefully. A security agent, network change, or automated containment action can disrupt physical processes if it is not tested for the relevant environment. Passive monitoring, segmented access paths, tightly controlled remote maintenance, offline recovery procedures, and joint exercises between IT and operational teams may be safer starting points than immediate full automation.

Table 3 Suggested implementation priorities by sector

Sector	Initial priorities	Potential AI-ZTA use
Public sector	Identity inventory, MFA, access reviews, secure backups, central logging	Anomaly detection across access logs; least-privilege access to sensitive records
Banking and payments	Privileged access management, transaction monitoring, supplier-risk review, incident exercises	Suspicious-session detection; step-up authentication for high-risk access
Telecommunications and critical infrastructure	Asset visibility, remote-access controls, network segmentation, tested recovery	Passive anomaly detection; tightly controlled access to operational systems
Universities and SMEs	Managed identity, endpoint protection, staff awareness, secure cloud configuration	Affordable managed detection; role-based access and risk alerts

Key implementation challenges

Skills and institutional capacity

AI security analytics and ZTA require expertise in identity management, cloud security, data engineering, incident response, and governance. The 2025 ISC2 Cybersecurity Workforce Study, which surveyed 16,029 practitioners and decision-makers worldwide, found that 95% reported at least one skills gap and that 88% had experienced at least one significant incident or operational issue linked to a skills shortage; AI (41%) and cloud security (36%) were the most frequently cited critical skills (ISC2, 2025). ISC2 did not publish a global workforce-gap estimate that year, arguing that respondents emphasised specific skills over headcount. Comparable Pakistan-specific workforce data were not identified for this paper, which is itself a research gap. Smaller organisations in Pakistan may lack dedicated security teams. Shared services, university–industry partnerships, professional training, and managed security services can help, and the National AI Policy’s capacity-building targets may support them (MoITT, 2025), but outsourcing must be accompanied by clear accountability and oversight.

Cost and legacy infrastructure

Licensing, log storage, network redesign, identity modernisation, and staff training can be expensive. Legacy applications may not support modern authentication or detailed logging, and vendor reporting links Pakistan's exposure to outdated systems and software vulnerabilities (Daily Times, 2025). A risk-based sequence, starting with critical assets, privileged accounts, backups, and externally exposed services, can reduce cost and implementation disruption. Because SIEM, EDR, and XDR deployments are log-intensive, organisations should also estimate storage and bandwidth costs before committing to a monitoring architecture, particularly where connectivity is constrained.

Data protection and sovereignty

Security analytics can process personal, financial, and operational information. Organisations should define what data is collected, why it is necessary, how long it is retained, who can access it, and whether it can be processed by a third party or outside Pakistan. Pakistan has long worked on a Personal Data Protection Bill, and legal commentary in 2026 continued to describe it as a draft, with no comprehensive enacted statute, no dedicated data protection authority, and no general breach-notification duty (Chambers and Partners, 2026; DLA Piper, n.d.); commentators have urged enactment (Khan, 2026). The draft bill has contemplated rules on cross-border transfer, localisation of critical personal data, and breach notification, but these remain proposals unless and until enacted. In this uncertain setting, organisations should apply privacy-by-design as a matter of policy, check data localisation, cross-border transfer, and retention decisions against applicable Pakistani laws, contracts, and sector-specific directions rather than assume one universal rule, and monitor legislative developments.

Third-party and supply-chain risk

Cloud providers, software suppliers, managed service providers, and outsourced support can become pathways into sensitive environments. The doubling of third-party involvement in breaches to 30% in the 2025 Verizon data illustrates the scale of this exposure globally (Verizon, 2025). Due diligence should cover access controls, breach notification, subcontractors, software updates, independent assurance, recovery commitments, and termination or migration plans. Vendor access should be time-limited, monitored, and restricted to the minimum required.

AI reliability, explainability, adversarial risk, and privacy

AI models may generate false alarms, miss unfamiliar threats, or be manipulated. Automated decisions can also affect legitimate users and business operations. Organisations should validate models using relevant data, maintain human review for consequential actions, monitor performance over time, and document how personal or sensitive data is handled (NIST, 2023). Procurement should ask vendors how models are protected against evasion and data poisoning,

what telemetry leaves the organisation, and how staff use of generative AI tools is governed, given the breach-cost penalty associated with shadow AI (IBM, 2025; Vassilev et al., 2025).

Coordination and incident readiness

Detection technology alone does not guarantee effective response. Organisations need named incident owners, clear escalation thresholds, contact lists, evidence-preservation procedures, tested backups, communications plans, and post-incident reviews. CERT and regulator reporting should be incorporated into exercises and operating procedures. The PKCERT remark that many organisations remain unaware of breaches underlines that detection capability, not only reporting discipline, is a limiting factor (Hanif, 2026).

Proposed strategic framework for Pakistan

This paper proposes a five-stage framework that can be adapted to an organisation's risk profile and resources. Each stage specifies activities, expected outputs, and indicative measures so that progress can be evaluated. The stages are cumulative but may overlap; the framework is aligned with the CISA maturity-stage logic and the NIST CSF 2.0 functions (CISA, 2023; NIST, 2024).

Stage 1: Establish governance and baseline

Assign executive ownership; identify critical services, data, users, devices, suppliers, and cloud dependencies; conduct a risk assessment; map applicable policy and regulatory obligations, including CERT reporting duties, SBP requirements where relevant, and the status of PISF 2026; and establish baseline indicators. The principal outputs are an asset and supplier register, a risk register, a data-flow inventory identifying personal and sensitive data, and an approved cybersecurity roadmap with budget. This stage should also decide how the organisation will meet the PKCERT security-operations expectation, whether in-house, through a managed provider, or through a sectoral arrangement.

Stage 2: Secure identities and essential assets

Enable multifactor authentication for privileged and remote access; remove dormant accounts; adopt least-privilege roles; secure administrative accounts; patch internet-facing systems; protect backups; and establish centralised, time-synchronised logging. Backups should be isolated or immutable and restoration should be tested, because recovery from backups is the control that determined the outcome of the 2026 government ransomware incident (Hanif, 2026). Given the prominence of credential abuse in global breach data (Verizon, 2025), this stage typically offers the highest security return for the lowest complexity.

Stage 3: Pilot AI-supported monitoring

Choose a bounded use case such as phishing triage, suspicious sign-in detection, or endpoint anomaly detection. Define success criteria in advance, compare alerts with analyst findings, measure false positives, and document model limitations before expanding. Pilots should record the data used, retention period, vendor access, and human-override procedure. Where possible, pilots should be designed to run in parallel with existing monitoring so that the organisation can compare detection and response times without increasing operational risk.

Stage 4: Introduce ZTA controls incrementally

Apply conditional access and device checks to selected systems, segment sensitive workloads, monitor third-party access, and introduce risk-based session controls. Test business continuity and emergency access before enforcing policies across critical operations. A useful sequence is to run new policies in report-only mode, review the exceptions they would have generated, remediate legacy dependencies, and only then enforce.

Stage 5: Evaluate, coordinate, and improve

Run tabletop exercises and technical tests; review incidents and model performance; update access policies and playbooks; coordinate with relevant CERTs and regulators; and publish internal performance reports for management oversight. Independent testing, such as red-team exercises and adversarial testing of AI components, should be scheduled at least annually for high-impact services.

Table 4 Proposed five-stage framework: activities, outputs, and indicators

Stage	Key activities	Principal outputs	Indicative measures
1. Governance and baseline	Executive ownership; asset, data, and supplier mapping; risk assessment; obligation mapping	Registers; roadmap and budget	Share of critical assets with owner and risk rating
2. Identities and essential assets	MFA; least privilege; dormant-account removal; patching; protected backups; central logging	Hardened identity layer; tested backups	Privileged accounts with MFA (%); backup restoration success
3. AI-supported monitoring pilot	Bounded use case; predefined success criteria; comparison with analyst findings	Pilot report; model documentation	Alert precision; false-positive rate; time to detect
4. Incremental ZTA	Conditional access; device checks; segmentation; vendor-access monitoring	Enforced policies; exception register	Excessive access removed; time to revoke access
5. Evaluate	Exercises; independent testing; CERT	Annual assurance	Exercise completion;

and improve	coordination; management reporting	report	share of incidents detected internally
-------------	------------------------------------	--------	---

Illustrative application scenarios

The following scenarios are hypothetical and are provided only to show how the framework could be applied. They are not case studies and do not report observed results.

A provincial government department. The department runs several citizen-facing applications and uses a mix of in-house and outsourced administrators. In Stage 1 it discovers that dozens of former contractors still have active accounts. Stage 2 removes them, introduces multifactor authentication for administrators and remote access, and centralises logs. In Stage 3 it pilots suspicious sign-in detection through a managed provider and finds, by comparing alerts with analyst review, that most false positives arise from shared workstations; it therefore adds device identity to the model. Stage 4 applies conditional access to the records system first. The measurable gains would be fewer dormant accounts, shorter time to revoke compromised access, and documented evidence for PKCERT reporting.

A mid-sized payment institution. The institution relies on a cloud-hosted platform and several technology vendors. Stage 1 maps its supplier access paths and SBP obligations. Stage 2 enforces privileged-access management and hardware-backed authentication for administrators. Stage 3 pilots behavioural analytics on administrative sessions, and Stage 4 introduces step-up authentication for high-risk operations such as changes to settlement instructions. Its principal risks are over-automation of account lockouts, which could interrupt payments, and vendor access to logs containing customer data; both are managed through human approval for high-impact actions and contractual limits on data use.

A university or SME. With no security team, the organisation adopts a managed identity service, enables multifactor authentication for all staff email accounts, and subscribes to managed detection for endpoints. Its AI use is entirely vendor-supplied, so governance focuses on contract terms, data location, and exit rights. Its indicators are modest but meaningful: percentage of accounts with multifactor authentication, time to restore from backup, and staff phishing-simulation results.

A telecommunications or energy operator. Because of OT dependencies, the operator begins with asset visibility and passive anomaly detection on network traffic, tightly controlled remote-maintenance sessions, and joint IT-OT exercises. Automated containment is deferred until it has been tested in a representative environment and approved by operations engineers.

Evaluation indicators and future research

A future empirical study should assess whether the proposed framework improves security outcomes in Pakistani organisations. Useful indicators include mean time to detect and contain

incidents; percentage of privileged accounts protected by multifactor authentication; proportion of critical assets with current owners and risk ratings; number of excessive or dormant accounts removed; alert precision and false-positive rate; time required to revoke compromised access; backup restoration success; percentage of critical suppliers assessed; and completion of incident-response exercises. These indicators should be interpreted alongside operational impacts, cost, user experience, and privacy risks. Because international studies link faster internal detection to lower breach costs (IBM, 2025), the share of incidents detected internally rather than reported by third parties is a particularly informative measure.

Future research could use multiple case studies across public administration, banking, telecommunications, higher education, and SMEs. Interviews with chief information security officers, IT managers, regulators, CERT personnel, and cloud service providers could identify barriers to adoption. A maturity model could compare governance, identity, data protection, monitoring, incident response, workforce capacity, and supplier management. Quantitative work could evaluate before-and-after metrics from controlled pilots, while ensuring that sensitive security information is protected. Cross-sector comparison would help determine which controls can be standardised nationally and which require sector-specific implementation. Four research questions are proposed:

- RQ1. What proportion of Pakistani organisations in each sector currently meets minimum identity, logging, and backup controls, and how does this vary by size and ownership?
- RQ2. To what extent do AI-supported monitoring pilots reduce time to detect and contain incidents relative to rule-based monitoring in the same organisation, and at what false-positive and cost levels?
- RQ3. Which organisational, legal, and workforce factors most strongly predict successful ZTA adoption in the public and financial sectors?
- RQ4. How do shared or managed security operation models perform compared with in-house centres for organisations that must meet the PKCERT security-operations expectation?

Limitations

This paper has several limitations. It is conceptual and does not test the framework empirically. Much of its Pakistan-specific evidence is drawn from press reports of statements and presentations, and from vendor telemetry, which should be replaced by primary documents and peer-reviewed studies in a journal version. Several instruments referred to, notably PISF 2026, the Personal Data Protection Bill, and the National Data Governance proposals, were described as draft or under review at the time of writing, and their status may have changed. International cost and effectiveness evidence comes mostly from high-income settings and from organisations that had experienced breaches, and may not generalise. Finally, the paper does not evaluate specific commercial products, and sector recommendations are necessarily general.

Discussion

The central argument is not that AI or ZTA alone will solve Pakistan's cybersecurity challenges. Rather, resilience depends on combining analytical capability with enforceable access controls, governance, skilled staff, reliable recovery, and coordination. AI is most valuable when it improves the quality and speed of security decisions; ZTA is most valuable when identity and access policies consistently limit the consequences of compromise. If underlying asset inventories, logs, identity records, and incident procedures are weak, advanced technology may merely automate unreliable processes.

Three tensions deserve attention. The first is between speed and assurance: the PKCERT deadline for security operations capability creates pressure to deploy quickly, whereas the evidence on AI risk and the dependence of ZTA on accurate inventories argue for staged deployment and testing. The second is between national coordination and organisational accountability: shared services and sectoral centres can fill capability gaps, but individual organisations remain responsible for their systems and suppliers. The third is between security analytics and privacy: monitoring improves detection but increases the volume of personal and operational data processed, in a setting where a comprehensive statutory data protection regime is still pending.

A Pakistan-oriented approach should be risk-based and proportionate. High-impact services need stronger assurance and more rigorous testing, while smaller organisations may begin with managed services and a small number of high-value controls. A phased approach also allows policymakers and organisations to build evidence about cost, effectiveness, and unintended consequences before scaling deployments. For policymakers, three actions follow: publish consolidated, comparable incident statistics with clear definitions; support shared or managed security-operation models for smaller institutions; and finalise data protection and AI governance rules so that organisations can adopt analytics with legal certainty.

Conclusion

Pakistan's growing dependence on digital systems makes cybersecurity an essential component of public service continuity, financial confidence, and critical infrastructure resilience. Reported figures, though not directly comparable, point to a substantial and evolving threat environment that includes malware, ransomware, and large-scale fraud. Cloud-based AI can support threat detection, prioritisation, and incident analysis, while Zero Trust Architecture can reduce implicit trust through continuous verification, least privilege, and segmentation. Together, these technologies offer a practical direction for improving security when introduced with clear governance and realistic expectations.

Pakistan already has relevant policy and institutional foundations, including the National Cyber Security Policy 2021, the CERT Rules 2023, the National AI Policy 2025, SBP technology-risk guidance, and the 2026 national directive on security operations (Business Recorder, 2026; Government of Pakistan, 2023; MoITT, 2021, 2025; SBP, 2025). The next challenge is to translate

these foundations into operational controls, skilled teams, tested response processes, responsible AI practices, and measurable assurance. This paper recommends a phased implementation programme beginning with governance, identity security, critical-asset visibility, and bounded AI pilots, followed by gradual ZTA expansion and continuous evaluation. The framework is conceptual and requires empirical testing across Pakistani sectors. Its value lies in offering a structured starting point for such research and for organisations seeking to improve cyber resilience.

References

- Arshad, K., Shaukat, J., & Bhatti, M. F. T. (2026). Cybersecurity and Public Trust in Autonomous Vehicles: Challenges, Mitigation Strategies, and Future Directions. *The Critical Review of Social Sciences Studies*, 4(1), 8399-8414.
- Bhatti, M. A. A., & Shaukat, J. (2025). Impact of Cybersecurity Risk Management Practices on Agile Project Performance: Mediating Role of Team Trust and Moderating Role of AI Adoption. *Research Consortium Archive*, 3(2), 1308-1335.
- Business Recorder. (2025, July 30). *Federal cabinet approves National AI Policy 2025*. <https://www.brecorder.com/news/amp/40375361>
- Business Recorder. (2026). *Cyber defence: Public, private entities told to set up 'SOCs' in 6 months*. <https://www.brecorder.com/news/40428085>
- Chambers and Partners. (2026). *Data protection & privacy 2026: Pakistan*. <https://practiceguides.chambers.com/practice-guides/data-protection-privacy-2026/pakistan>
- Chandramouli, R., & Butcher, Z. (2023). *A zero trust architecture model for access control in cloud-native applications in multi-location environments* (NIST Special Publication 800-207A). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-207A>
- Cybersecurity and Infrastructure Security Agency. (2023). *Zero trust maturity model* (Version 2.0). U.S. Department of Homeland Security. <https://www.cisa.gov/zero-trust-maturity-model>
- Daily Times. (2025, November 14). *Pakistan records over 5.3 million cyberattacks in 2025*. <https://dailytimes.com.pk/1400414/pakistan-records-over-5-3-million-cyberattacks-in-2025/>
- DataReportal. (2025). *Digital 2026: Pakistan*. Kepios. <https://datareportal.com/reports/digital-2026-pakistan>
- Dawn. (2025, July 30). *Federal cabinet approves National AI Policy 2025*. <https://www.dawn.com/news/1927634/federal-cabinet-approves-national-ai-policy-2025>
- DLA Piper. (n.d.). *Data protection laws of the world: Pakistan*. Retrieved October 20, 2026, from <https://www.dlapiperdataprotection.com/index.html?t=law&c=PK>

- Government of Pakistan. (2023, October 3). *Computer Emergency Response Team Rules, 2023* [Gazette of Pakistan, Extraordinary]. <https://pkcert.gov.pk/uploads/2023/10/GAZETTE-CERT-Rules-2023.pdf>
- Hanif, U. (2026, July 2). Pakistan faces 253 cyberattacks in 2026; govt issues COC deadline for organisations. *The Express Tribune*. <https://tribune.com.pk/story/2616013/us-iran-deal-role-spurs-cyber-threats>
- IBM. (2025). *Cost of a data breach report 2025*. <https://www.ibm.com/reports/data-breach>
- ISC2. (2025). *2025 ISC2 cybersecurity workforce study*. <https://www.isc2.org/insights/2025/12/2025-isc2-cybersecurity-workforce-study>
- Khan, S. (2026, August 24). Pakistan can no longer wait for personal data protection law. *The Express Tribune*. <https://tribune.com.pk/story/2625470/pakistan-can-no-longer-wait-for-personal-data-protection-law>
- Kolawole, I. (2025). Leveraging cloud-based AI and zero trust architecture to enhance U.S. cybersecurity and counteract foreign threats. *World Journal of Advanced Research and Reviews*, 25(3), 6–25. <https://doi.org/10.30574/wjarr.2025.25.3.0635>
- Ministry of Information Technology and Telecommunication. (2021). *National Cyber Security Policy 2021*. Government of Pakistan. <https://pkcert.gov.pk/uploads/2023/06/2021-National-Cyber-Security-Policy-Final.pdf>
- Ministry of Information Technology and Telecommunication. (2025). *National Artificial Intelligence Policy 2025*. Government of Pakistan. <https://issi.org.pk/issue-brief-on-pakistans-national-ai-policy-2025-developing-ai-ecosystem-and-ethical-data-governance/>
- National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AIRMF1.0)*(NIST AI 100-1). <https://doi.org/10.6028/NIST.AI.100-1>
- National Institute of Standards and Technology. (2024). *The NIST cybersecurity framework (CSF) 2.0*(NIST CSWP 29). <https://doi.org/10.6028/NIST.CSWP.29>
- Pakistan Telecommunication Authority. (2024, September 14). *Pakistan ranks among top countries in Global Cybersecurity Index 2024, marking significant progress*. <https://pta.gov.pk/category/pakistan-ranks-among-top-countries-in-global-cybersecurity-index-2024,-marking-significant-progress-1485299784-2024-09-23>
- Profit. (2025, October 23). *Cybercrime surges 35% in Pakistan amid rising WhatsApp hacking and digital fraud cases*. <https://profit.pakistantoday.com.pk/2025/10/23/cybercrime-surges-35-in-pakistan-amid-rising-whatsapp-hacking-and-digital-fraud-cases/>
- ProPakistani. (2025, November 14). *Kaspersky reveals shocking details about cyber attacks on Pakistanis in 2025*. <https://propakistani.pk/2025/11/14/kaspersky-reveals-shocking-details-about-cyber-attacks-on-pakistanis-in-2025/>
- Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero trust architecture* (NIST Special Publication 800-207). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-207>

- State Bank of Pakistan. (n.d). *Banking regulation and supervision: Cybersecurity and technology risk supervision*. <https://www.sbp.org.pk/our-operations/banking-regulation-and-supervision>
- State Bank of Pakistan. (2025, October 3). *Technology risk management framework for payment institutions* (PSPOD Circular No. 04 of 2025). <https://www.sbp.org.pk/circulars/pspod-circular-no-04-of-2025>
- Umar, M., Shaukat, J., & Bhatti, M. F. T. (2026). Artificial Intelligence in Cybersecurity: Threat Detection, Defense, and the Governance of a Dual-Use Technology. *Research Consortium Archive*, 4(2), 4951-4969.
- Vassilev, A., Oprea, A., Fordyce, A., & Andersen, H. (2024). Adversarial machine learning: A taxonomy and terminology of attacks and mitigations.
- Verizon. (2025). *2025 data breach investigations report*. <https://www.verizon.com/business/resources/reports/dbir/>