© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723189

Artificial Intelligence for Autonomous Database Management in Saudi Arabia: Predictive Performance Monitoring and Intelligent Incident Prevention for Vision 2030

MOHAMMED AFZAL SHAREEF

Abstract- ***Saudi Arabia's Vision 2030 is accelerating data-intensive digital government, cloud adoption, artificial intelligence (AI), cybersecurity, and highly available national platforms. As mission-critical database estates grow, conventional administration based on static thresholds, manual diagnosis, and reactive incident handling becomes difficult to scale. This paper examines how AI can support autonomous database management in Saudi Arabia through predictive performance monitoring, anomaly detection, root-cause analysis, and guardrailed incident prevention. A structured narrative review of 30 core sources published mainly between 2020 and 2026 synthesizes peer-reviewed database and AIOps research, selected industry documentation, and official Saudi digital and AI policy. The review shows that effective autonomy requires more than anomaly detection: heterogeneous telemetry must be transformed into contextual features, models must adapt to workload drift, anomalies must be linked to probable causes, and remediation must be constrained by explainability, policy, rollback, security, and human oversight. Building on these findings, the paper proposes a six-layer Saudi Autonomous Database Operations Framework spanning observability, operational data engineering, AI analytics, contextual diagnosis, decision orchestration, and governance. A five-stage maturity model guides progression from observable operations to policy-governed closed-loop automation. The framework aligns with Saudi priorities for data and AI adoption, advanced digital infrastructure, reliability, cybersecurity, responsible AI, and efficient digital services. The paper concludes that AI can strengthen database resilience when autonomy is introduced incrementally and governed by trustworthy telemetry, validated models, architecture-aware diagnosis, and controlled automation.***

Keywords: ***autonomous database management; artificial intelligence; AIOps; anomaly detection; predictive monitoring; root-cause analysis; database reliability; Saudi Vision 2030; digital transformation; incident prevention***

I. INTRODUCTION

Database systems are foundational to modern digital economies because they support financial transactions, government services, healthcare platforms, identity systems, analytics, enterprise resource planning, and increasingly AI-enabled applications. In Saudi Arabia, this role is becoming more important as Vision 2030 moves the Kingdom from capability building toward scaled digital and knowledge-economy impact. The Vision 2030 Annual Report 2025 describes sustained investment in digital infrastructure, artificial intelligence, cybersecurity, innovation, and digital services, while the National Transformation Program reports continued progress in digital government maturity and national digital infrastructure. These developments create a corresponding requirement for reliable, secure, and continuously available data platforms (Saudi Vision 2030, 2026; National Transformation Program, 2026).

Traditional database operations depend heavily on DBAs interpreting dashboards, alarms, wait events, SQL plans, replication states, and user incidents. This model becomes increasingly reactive as estates expand across Oracle, PostgreSQL, Microsoft SQL Server, cloud databases, warehouses, and hybrid infrastructure. Static thresholds may flood operators during predictable peaks while missing multivariate deterioration. A CPU spike, for example, can reflect inefficient SQL, lock contention, plan regression, statistics drift, storage latency, memory pressure, or application change. The central problem is therefore not simply whether a metric is abnormal, but whether the deviation is meaningful, what caused it, how soon it may threaten a service objective, and which response is safe.

Research on AI for databases and Artificial Intelligence for IT Operations (AIOps) offers a route

toward more autonomous operations. AI4DB research has explored learned configuration, query optimization, index selection, self-tuning, workload prediction, and database security (Zhou et al., 2022; Zou et al., 2022). AIOps research has addressed anomaly detection, failure prediction, root-cause analysis, and automated remediation across complex IT environments (Notaro et al., 2021; Diaz-De-Arcaya et al., 2024). More recent database-specific research shows growing emphasis on KPI-based performance anomaly detection, learned components inside database engines, and production-scale root-cause diagnosis (Gao et al., 2025; Dritsas and Trigka, 2026; Liu et al., 2022).

For Saudi Arabia, the relevance is strategic as well as technical. The Saudi Data and Artificial Intelligence Authority (SDAIA) is the national reference for data and AI and has explicitly connected AI adoption, advanced digital infrastructure, reliability, cybersecurity, government cloud services, governance, and responsible AI with national ambitions (SDAIA, n.d.-a; SDAIA, n.d.-b). The AI Adoption Framework provides a structured approach to organizational AI adoption, while the AI Ethics Principles emphasize reliability, safety, transparency, accountability, privacy, and risk management (SDAIA, 2024; SDAIA, 2023). These principles imply that an autonomous database platform deployed in government or critical enterprise contexts should not be defined as a system that acts without control. Rather, autonomy should mean progressively reducing repetitive manual work while improving evidence-based decisions, with human oversight and policy controls proportional to operational risk.

This paper addresses three questions. First, which AI methods are most relevant to predictive database monitoring and incident prevention? Second, how should these methods be integrated into an operational architecture that moves from detection to diagnosis and safe action? Third, how can such an architecture align with Saudi Vision 2030 priorities for digital transformation, data and AI leadership, resilience, cybersecurity, and responsible governance? The contribution is a structured synthesis and a proposed Saudi Autonomous Database Operations Framework designed for heterogeneous enterprise environments rather than a single database product. The paper deliberately avoids presenting fabricated production experiments; instead, it develops a conceptual, implementation-oriented framework grounded in the literature and current Saudi policy direction.

II. LITERATURE BACKGROUND AND RESEARCH CONTEXT

2.1 From automated administration to AI-enabled database autonomy

Database automation predates current AI systems. Conventional DBMS products already automate backup scheduling, statistics maintenance, plan selection, memory management, failover, and diagnostic collection to varying degrees. The recent shift is the use of machine learning to learn from changing workloads and operational history rather than relying only on fixed rules. Zhou et al. (2022) describe AI4DB as the use of AI to improve configuration tuning, optimization, indexing, advising, and security, while Zou et al. (2022) characterize learnable databases as systems that adapt to workload and data characteristics. Deep learning has also been studied across query execution, including cardinality estimation, indexing, query optimization, and tuning (Milicevic and Babovic, 2024). A 2026 survey by Dritsas and Trigka further emphasizes that learned components change not only performance but also system-level concerns such as retraining, latency, robustness, and explainability.

Commercial autonomous database offerings demonstrate the operational direction of this research. Oracle Autonomous AI Database automates provisioning, backup, patching, scaling, tuning, and other lifecycle tasks, while Oracle's Autonomous Health Framework combines continuous monitoring, machine-learning-assisted diagnostics, and corrective guidance intended to reduce reaction time (Oracle, 2026a; Oracle, 2026b). These systems show that autonomy is most useful when it integrates telemetry, domain knowledge, and operational actions rather than treating machine learning as an isolated prediction service.

2.2 Predictive monitoring and anomaly detection

Database performance monitoring produces multivariate time-series data. Typical signals include CPU consumption, memory pressure, physical and logical I/O, transaction rate, query latency, throughput, active sessions, wait classes, lock

duration, buffer cache behavior, redo or transaction-log generation, connection pool saturation, replication lag, failed jobs, backup status, storage latency, error logs, and operating-system metrics. Static thresholds remain useful for clear safety limits, but they struggle with dynamic baselines and correlated behavior.

Deep anomaly-detection research provides techniques for learning normal behavior and identifying deviations in high-dimensional data. Pang et al. (2021) classify deep anomaly detection into major methodological families and highlight the value of representation learning for complex data. USAD uses adversarially trained autoencoders for unsupervised multivariate time-series anomaly detection and was motivated by the scale and complexity of IT monitoring data (Audibert et al., 2020). LogBERT applies self-supervised transformer learning to system logs, learning normal event-sequence patterns and identifying deviations (Guo et al., 2021). Benchmarking research also warns that algorithm choice and evaluation matter: Schmidl et al. (2022) evaluated a large set of time-series anomaly-detection algorithms and found substantial variation across datasets and anomaly types, while Cao et al. (2025) show that streaming environments introduce concept drift and require model adaptation.

For databases specifically, Gao et al. (2025) review KPI-based performance anomaly detection and note that database anomalies require domain-specific interpretation and root-cause analysis. This is important because a model that labels an observation as anomalous but cannot identify the affected database object, SQL workload, resource, or dependency provides limited operational value. Accordingly, the goal of autonomous database monitoring should be early detection with contextual localization rather than a single anomaly score.

2.3 Root-cause analysis and incident prevention

AIOps research commonly separates failure management into detection, diagnosis, prediction, and remediation, a lifecycle also emphasized in the AIOps workshop synthesis (Bogatinovski et al., 2021). Notaro et al. (2021) survey AIOps methods for failure management and show that interventions differ according to whether they occur before, during, or after a failure. Diaz-De-Arcaya et al. (2024) similarly identify AIOps as a growing operational paradigm but note gaps in lifecycle integration and deployment maturity. Predictive incident research emphasizes the value of using historical incident and operational data to estimate risk before service impact (Ahmed et al., 2023).

In database environments, causal diagnosis is especially important because symptoms propagate across layers. Liu et al. (2022) present PinSQL, a production-oriented system for cloud databases that collects performance metrics and query logs, detects anomalies, identifies root-cause SQL templates, and recommends or executes repair actions. Their work demonstrates the practical difference between correlated high-impact SQL and true root-cause SQL. This distinction is central to safe automation: terminating a highly active query that is only a symptom can worsen the incident, while addressing the actual root-cause query or resource bottleneck can restore stability.

Explainability also becomes operationally necessary. Ouared et al. (2023) argue that AI-assisted database tuning should keep the DBA in the loop by providing transparent reasoning for tuning actions. Li et al. (2024) survey explainable anomaly detection and show that actionable explanations are a distinct problem beyond prediction accuracy. For high-risk database actions such as failover, parameter change, session termination, plan forcing, or resource scaling, explanation should therefore be treated as a control requirement rather than an optional dashboard feature.

2.4 Saudi digital and AI policy context

The National Strategy for Data and AI positions data and AI as instruments for economic and social value and national leadership in data-driven economies (SDAIA, n.d.-a). SDAIA's current strategic objectives include AI adoption, advanced digital infrastructure, reliability, cybersecurity, and government cloud services (SDAIA, n.d.-b). The 2024 AI Adoption Framework encourages structured, responsible adoption across sectors, while national AI ethics guidance requires ongoing risk management, performance monitoring, interpretability, privacy, and governance (SDAIA, 2024; SDAIA, 2023). In 2026, SDAIA also published a National AI Risk Management Framework, reinforcing lifecycle-based identification, assessment, treatment, and monitoring of AI risks (SDAIA, 2026).

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723189

These policies are directly relevant to autonomous database management. Database telemetry can contain sensitive operational and business information; models can make incorrect or biased decisions due to drift; and remediation actions can affect nationally important services. Therefore, a Saudi implementation should combine predictive intelligence with cybersecurity, privacy, approval policy, traceability, and rollback. The Digital Government Authority's guidance on cloud integration further links cloud adoption with government efficiency, service quality, information security, and digital transformation (Digital Government Authority, 2025). The technical objective is consequently not maximum automation at any cost, but measurable improvement in reliability and efficiency within a governed operating model.

III. METHODOLOGY

This study uses a structured narrative review and conceptual design method. The purpose is to synthesize recent research and policy evidence into an implementation framework rather than conduct a statistical meta-analysis. Searches were performed up to 3 September 2026 using combinations of terms including "autonomous database", "AI4DB", "AIOps", "database anomaly detection", "predictive monitoring", "root cause SQL", "time-series anomaly detection", "explainable anomaly detection", "Saudi AI adoption", "Saudi Vision 2030 digital transformation", and "data and AI governance". Priority was given to peer-reviewed journal and conference sources from ACM, IEEE, Elsevier, Springer, and VLDB, together with official Saudi government sources and selected vendor technical documentation relevant to deployed autonomous database capabilities.

The final evidence base used in the paper contains 30 core sources: 19 peer-reviewed journal or conference studies, one research white paper, two industry technical sources, and eight official Saudi strategy, governance, or digital-transformation sources. Most research sources were published from 2020 onward; older foundational material was not required because recent surveys already synthesize earlier work. Inclusion criteria were direct relevance to at least one of five themes: AI-enabled database optimization, anomaly detection, AIOps/failure management, database root-cause diagnosis, or Saudi digital/AI governance. Sources that made unverifiable performance claims or lacked sufficient bibliographic information were excluded.

The synthesis followed four analytical steps. First, operational database tasks were mapped to AI capabilities. Second, technical limitations such as false positives, drift, explainability, and unsafe automation were identified. Third, these capabilities and controls were assembled into a layered architecture. Fourth, the architecture was mapped to Saudi national priorities. Because no production dataset was supplied for this study, the proposed framework and maturity model are conceptual; no empirical performance improvement is claimed. This distinction is important for research integrity and for later journal review.

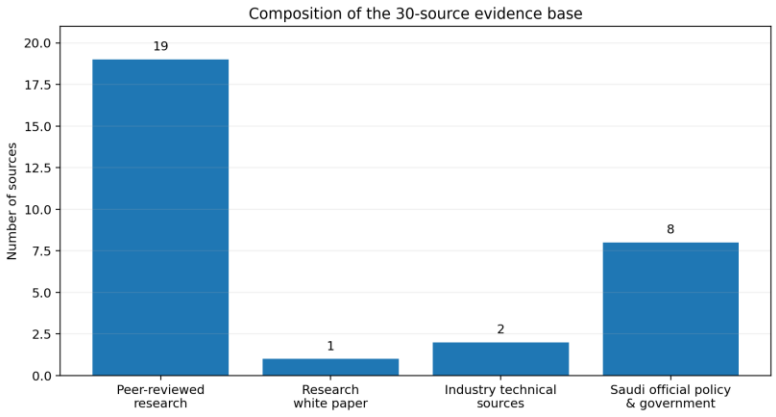

Figure 1. Composition of the review evidence base (n = 30). Source: author synthesis.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723189

IV. PROPOSED SAUDI AUTONOMOUS DATABASE OPERATIONS FRAMEWORK

The proposed framework contains six interacting layers: observability, operational data engineering, AI analytics, causal diagnosis, decision orchestration, and governance. A feedback loop connects executed actions to subsequent monitoring so that the system can learn from outcomes without allowing uncontrolled self-modification.

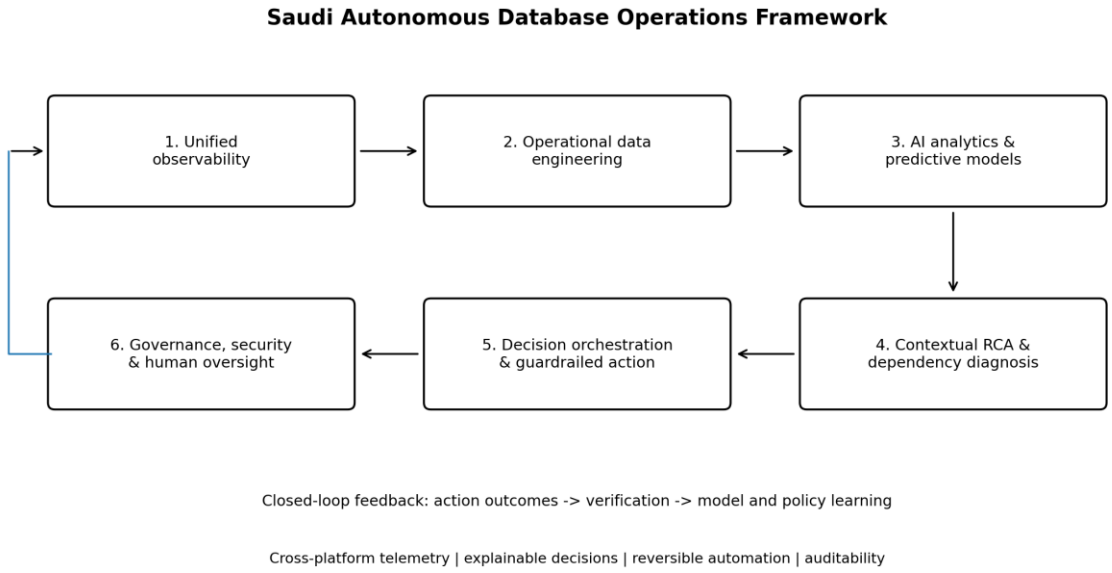

Figure 2. Six-layer Saudi Autonomous Database Operations Framework. Source: author conceptualization based on reviewed literature and Saudi governance requirements.

4.1 Layer 1: unified database observability

Autonomous operations begin with high-quality telemetry. A heterogeneous Saudi enterprise may operate Oracle RAC and Data Guard, PostgreSQL, Microsoft SQL Server, cloud-managed databases, ERP databases, analytics platforms, and replication services. Each technology exposes different internal metrics, yet the operational questions are similar: Is performance degrading? Is capacity sufficient? Is availability at risk? Has a change introduced instability? Is the problem local to one database, one SQL workload, or a dependent infrastructure component?

The observability layer should normalize metrics, logs, traces, wait events, configuration state, SQL statistics, operating-system signals, storage/network measurements, replication state, backup events, and change records into a time-aligned telemetry model. Context must be preserved: a CPU spike during an expected reporting batch differs from the same spike on an otherwise idle database, and replication lag during planned bulk loading differs from lag under normal transaction volume. Patching, schema deployment, parameter changes, statistics collection, and application releases should be recorded because they frequently explain performance shifts.

Telemetry quality controls must precede machine learning because missing values, clock skew, duplicate alerts, unstable metric names, and monitoring gaps can create false anomalies. Telemetry should also carry service criticality, ownership, recovery objectives, data classification, and business-calendar context so models can distinguish technical deviation from business significance.

4.2 Layer 2: operational data engineering and feature construction

Raw observability data must be converted into features that represent database behavior. Useful features

include rolling means and percentiles, rate of change, seasonality residuals, wait-class ratios, I/O latency distributions, buffer and memory pressure indicators, query-shape fingerprints, plan changes, lock graphs, replication-lag trends, error frequencies, connection churn, and relationships between workload and resource consumption. Feature windows should be created at multiple scales because incidents may develop over seconds, minutes, or hours.

Cross-platform normalization is especially valuable. Rather than forcing all products into identical low-level metrics, the framework can define semantic features such as "transaction pressure", "read I/O stress", "write-log pressure", "concurrency contention", "query regression", "replication risk", and "capacity saturation". Product-specific measurements are then mapped to these semantic categories. This creates a common AI layer while preserving vendor-specific details for diagnosis.

Operational data engineering also needs labels where available. Incident tickets, post-incident reviews, known maintenance periods, root-cause categories, and successful remediation records can be linked to telemetry. However, fully labeled datasets are rare; unsupervised and semi-supervised approaches remain important because most production time represents normal behavior and novel failures may not have historical labels. This supports hybrid modelling rather than dependence on a single supervised classifier.

4.3 Layer 3: predictive analytics and anomaly detection

The AI analytics layer should combine multiple model families. Simple statistical baselines and robust thresholds remain useful for transparent safeguards. Isolation-based and density-based methods can identify outliers with relatively low complexity. Autoencoders can learn multivariate normal behavior and flag high reconstruction error, while recurrent or transformer models can capture temporal dependencies in metrics and logs. Forecasting models can predict future capacity, latency, or replication lag, allowing the system to estimate when a service-level threshold may be breached rather than waiting for a present-time alarm.

A practical platform should use ensemble reasoning so that a potential incident is supported by multiple signals, such as a workload forecast, anomaly score, and business-criticality rule. Models should report confidence and support adaptive baselines. Streaming-anomaly research shows that concept drift degrades accuracy as distributions change (Cao et al., 2025); databases are exposed to drift through releases, seasonality, schema evolution, data growth, hardware change, and cloud scaling. Retraining or recalibration should therefore be governed by drift detection and validation.

Predictive monitoring is most valuable when it estimates operational lead time. Instead of merely predicting "anomaly/no anomaly", the system can estimate probability of saturation or failure within a future horizon. Examples include forecasted tablespace exhaustion, memory pressure, increasing replication lag, rising lock duration, approaching connection limits, abnormal redo growth, backup-window overrun, or a trend toward SLA latency violation. This converts AI output into an actionable risk signal.

Table 1. AI techniques applicable to predictive and autonomous database operations.

Technique / model family	Typical database inputs	Primary operational use	Strength	Key limitation
Dynamic/statistical baselines	KPIs, seasonality, workload calendars	Context-aware thresholding	Transparent; low operational cost	Limited for complex multivariate interactions
Isolation / density methods	Multivariate KPIs	Outlier detection	Works with limited labels	Sensitive to feature scaling and drift
Autoencoders	Multivariate time-series telemetry	Learn normal state; detect deviations	Captures nonlinear relationships	Threshold selection and explainability

Technique / model family	Typical database inputs	Primary operational use	Strength	Key limitation
LSTM / Transformer models	Metrics, event sequences, logs	Forecasting and sequence anomalies	Models temporal dependence	Compute, retraining and concept drift
Supervised risk models	Historical incidents + engineered features	Failure / SLA risk prediction	Directly optimized for known incidents	Requires reliable labels; rare-event imbalance
Graph / causal models	Topology, SQL, change and dependency graphs	Root-cause ranking	Architecture-aware diagnosis	Graph quality and causal ambiguity
Explainable AI	Features + model outputs	Operator-facing evidence	Supports trust and governance	Explanations may not equal causality
Bounded reinforcement learning	State, actions, rewards, policy constraints	Optimization / remediation selection	Potential closed-loop adaptation	Unsafe without strong constraints and simulation

Source: author synthesis from Pang et al. (2021), Audibert et al. (2020), Guo et al. (2021), Gao et al. (2025), Liu et al. (2022), Li et al. (2024) and related reviewed literature.

4.4 Layer 4: contextual root-cause and dependency diagnosis

An anomaly becomes operationally useful only when it can be connected to likely causes. Root-cause analysis should combine temporal correlation with topology and domain knowledge. The system should know relationships among applications, services, database instances, schemas, SQL templates, storage systems, hosts, clusters, replication partners, and network dependencies. It should also know which changes occurred before the anomaly.

Diagnosis can follow a structured sequence: identify the affected service and anomaly window; determine which semantic feature groups changed first; correlate those changes with SQL workload, configuration events, waits, locks, I/O, and dependencies; rank candidate causes with causal, probabilistic, graph-based, or supervised models; and produce an operational explanation. Rather than returning only "anomaly score 0.93," the system should identify the evidence, likely cause, unaffected components, and confidence supporting the diagnosis.

PinSQL demonstrates how production cloud diagnosis can distinguish root-cause SQL from merely high-impact SQL (Liu et al., 2022). Similar reasoning can be extended to lock chains, replication, memory pressure, or storage bottlenecks. Explainability research supports this approach because operators need to understand feature contribution and causal sequence before accepting disruptive actions (Li et al., 2024; Ouared et al., 2023).

4.5 Layer 5: decision orchestration and guardrailed remediation

The decision layer converts diagnoses into responses. Actions should be classified by reversibility and risk. Low-risk actions may be fully automated: collect additional diagnostics, increase monitoring frequency, open an incident, enrich a ticket, notify the responsible team, or run a non-invasive health check. Medium-risk actions may be automated under predefined policy: scale a cloud resource within a bounded range, pause a non-critical batch job, refresh a monitoring agent, or switch to a pre-approved resource group. High-risk actions should normally require approval or strict multi-signal validation: database failover, killing sessions, forcing execution plans, changing optimizer parameters, modifying memory allocation, disabling jobs, or changing replication configuration.

Every automated action should have five controls. First, a precondition check verifies that the diagnosed state still exists. Second, policy authorization confirms that the action is permitted for the service and time

window. Third, a rollback path is defined. Fourth, post-action verification tests whether the intended metric improved and whether secondary risk emerged. Fifth, a complete audit record stores the evidence, model version, explanation, action, approver if applicable, and outcome. This structure converts automation from a script library into a governed control system.

4.6 Layer 6: governance, security, and human oversight

Governance spans all layers rather than operating as a final review step. The Saudi AI Adoption Framework and AI Ethics Principles support lifecycle governance, responsible use, transparency, and risk management (SDAIA, 2024; SDAIA, 2023). For autonomous database operations, this means the platform should manage model inventories, training data lineage, access control, approval policy, monitoring of model drift, adversarial or poisoned telemetry risk, segregation of duties, privacy, and cybersecurity.

Human oversight is particularly important during early maturity. DBAs and site reliability engineers should be able to accept, reject, or modify recommendations, and those decisions can become feedback for future model improvement. Explainability must be understandable to practitioners rather than expressed only in model-specific mathematics. The aim is a collaborative operating model in which AI performs high-volume correlation and prediction while experienced engineers retain authority over critical decisions until sufficient evidence supports higher autonomy.

V. PREDICTIVE PERFORMANCE MONITORING AND INTELLIGENT INCIDENT PREVENTION

The proposed framework supports five major operational use cases.

First, dynamic performance baselining replaces fixed thresholds with time-aware expectations. A model can learn normal ranges by hour, weekday, business cycle, and workload type. Deviations are then assessed against expected context. This reduces unnecessary alerts while improving sensitivity to subtle changes. Deep anomaly detection is useful when multiple weak signals combine into a meaningful pattern that would not cross individual thresholds (Pang et al., 2021; Gao et al., 2025).

Second, capacity and saturation forecasting estimates future resource risk. Time-series models can project growth in storage, sessions, memory, I/O queues, transaction logs, and replication lag. Predictive analytics research shows the broader value of historical pattern learning for forecasting future states (Jamarani et al., 2024). In database operations, forecast outputs should be linked to time-to-threshold so teams can act before user impact. A forecast that storage will reach 90% in 18 days is more useful than a warning generated only after the threshold is exceeded.

Third, query and plan regression detection identifies abnormal SQL behavior. Query fingerprints can be monitored for changes in latency, CPU per execution, logical reads, physical reads, rows processed, plan hash, and concurrency. A model can compare current performance with workload-normalized history. When regression is detected, the diagnosis layer can check recent statistics, indexes, schema changes, application releases, and parameter modifications. Deep learning research on query execution and AI4DB demonstrates growing opportunities for learned optimization, but production adoption should retain interpretable fallback mechanisms because query behavior can change abruptly (Milicevic and Babovic, 2024; Dritsas and Trigka, 2026).

Fourth, failure precursor detection searches for weak signals that precede known incidents. Historical incident windows can be aligned with telemetry to identify recurring precursors such as rising wait times, repeated connection resets, increasing log errors, unstable replication, or a combination of resource pressure and workload growth. AIOps research frames this as intervention before failure rather than diagnosis after failure (Notaro et al., 2021). The value should be evaluated through lead time, precision, and avoided service impact, not only classification accuracy.

Fifth, intelligent incident prevention combines prediction with safe action. Consider a database where active sessions are rising, lock waits are accelerating, and a new SQL template has appeared after an application deployment. The system can identify the pattern, rank the new SQL as a probable cause, collect its plan and blocking graph, notify the application

owner, and—if approved policy permits—throttle the relevant workload or pause a non-critical job. If the situation continues to deteriorate, a higher-severity workflow can be activated. This staged response is safer than immediately killing sessions based on a single alert.

Evaluation should therefore use both machine-learning and operational metrics. Precision, recall, F1-score, AUC-PR, and false-positive rate are necessary, but they are insufficient. Operational evaluation should include mean time to detect, diagnosis accuracy, time-to-root-cause, useful warning lead time, alert reduction, percentage of incidents prevented, change-failure rate from automated actions, rollback success, mean time to recover, and service-level impact. Schmidl et al. (2022) and Wenig et al. (2022) emphasize the importance of rigorous anomaly-detection evaluation, while database-specific deployment requires extending metrics to action quality and business impact.

Table 2. Recommended evaluation metrics for an autonomous database operations pilot.

Evaluation dimension	Example metrics	Why it matters operationally
Detection quality	Precision, recall, F1, AUC-PR, false-positive rate	Measures whether alerts are both sensitive and actionable
Timeliness	MTTD, warning lead time, time-to-threshold	Shows whether prediction creates enough time for prevention
Diagnosis	Root-cause top-k accuracy, time-to-root-cause	Separates useful diagnosis from generic anomaly alarms
Service reliability	MTTR, downtime, SLA/SLO impact	Connects model output to user-facing continuity
Automation safety	Change-failure rate, rollback success, remediation success	Tests whether autonomous actions reduce rather than amplify risk
Operational efficiency	Alert reduction, repetitive tickets avoided, utilization	Quantifies reduced toil and better infrastructure use
Model governance	Drift rate, override rate, explanation coverage, audit completeness	Supports responsible AI lifecycle control

Source: author synthesis; anomaly-detection benchmarking principles informed by Schmidl et al. (2022) and Wenig et al. (2022).

VI. ALIGNMENT WITH SAUDI VISION 2030

The proposed framework aligns with Saudi Vision 2030 in five practical dimensions.

Digital government reliability. Vision 2030 and the National Transformation Program emphasize advanced digital services and government technology maturity. Database availability is a prerequisite for reliable digital identity, citizen services, financial systems, licensing, healthcare, and public-sector applications. Predictive monitoring can support continuity by identifying degradation before it

becomes an outage, while automated diagnostics can reduce recovery time (Saudi Vision 2030, 2026; National Transformation Program, 2026).

Data and AI leadership. SDAIA's National Strategy for Data and AI seeks to create economic and social value from data and position Saudi Arabia among leading data-driven economies. Applying AI to the operational layer of data platforms extends AI adoption beyond front-end applications into the infrastructure that supports them. This is strategically significant because trustworthy AI services depend on reliable data stores, consistent performance, and secure data movement (SDAIA, n.d.-a).

Advanced digital infrastructure and cloud. SDAIA's strategic objectives include advanced digital infrastructure and high-standard government cloud services, and DGA guidance links cloud adoption with efficiency, integration, service quality, and information security (SDAIA, n.d.-b; Digital Government Authority, 2025). Autonomous database operations can support hybrid and cloud architectures by normalizing telemetry across platforms, predicting capacity requirements, and coordinating scalable responses without removing governance.

Cybersecurity and resilience. Availability, integrity, and confidentiality are linked. A database anomaly may represent performance degradation, misconfiguration, abuse, or security activity. Autonomous monitoring should therefore integrate with security operations while maintaining separation between performance and security diagnoses. AI can improve pattern recognition, including anomaly-oriented security monitoring (Ruffo et al., 2024), but cybersecurity controls must protect the monitoring pipeline itself. An attacker who can manipulate telemetry, model inputs, or remediation policy could convert an autonomous platform into an attack path. Saudi national priorities for cybersecurity and resilience make secure-by-design implementation essential.

Responsible and efficient AI adoption. SDAIA's AI Adoption Framework, AI Ethics Principles, and 2026 AI risk-management direction emphasize governance, reliability, transparency, privacy, and monitoring across the AI lifecycle (SDAIA, 2024; SDAIA, 2023; SDAIA, 2026). The proposed framework incorporates these requirements through model governance, explainability, approval controls, audit records, rollback, and human oversight. In this sense, responsible AI is not separate from performance engineering; it is the mechanism that allows autonomous operations to scale safely.

The national value proposition should be measured in operational outcomes rather than AI adoption alone. Relevant indicators include reduced unplanned downtime, lower mean time to diagnosis, fewer repetitive manual incidents, improved infrastructure utilization, earlier capacity planning, reduced alert fatigue, improved compliance evidence, and a growing base of Saudi professionals skilled in both database engineering and AI operations. This supports the broader Vision 2030 objective of building advanced technical capabilities while improving the efficiency and reliability of digital services.

Table 3. Alignment of autonomous database management with Saudi Vision 2030 and national digital priorities.

Vision 2030 / national theme	Autonomous database contribution	Illustrative KPI
Reliable digital government	Earlier detection, diagnosis and controlled recovery	Downtime; MTTD; MTTR; SLO attainment
Data and AI leadership	AI embedded in the operational data layer	% of database estate using AI-assisted operations
Advanced digital infrastructure & cloud	Cross-platform observability, forecasting and bounded scaling	Capacity headroom; cloud utilization; saturation events
Cybersecurity & resilience	Telemetry integrity, anomaly correlation, controlled remediation	Security-linked incident rate; recovery objective attainment
Responsible AI	Explainability, human oversight, audit and model governance	Override rate; drift incidents; audit coverage

Vision 2030 / national theme	Autonomous database contribution	Illustrative KPI
Human capability development	DBA/SRE transition to higher-value engineering and AI oversight	Training coverage; advanced operations competency

Source: author mapping based on Saudi Vision 2030 (2026), National Transformation Program (2026), SDAIA strategy and governance documents, and DGA cloud guidance.

VII. IMPLEMENTATION ROADMAP AND DISCUSSION

A realistic adoption roadmap should progress through controlled maturity stages. Stage 1: observable operations establishes reliable telemetry, service ownership, data quality, incident taxonomy, and change correlation. Organizations should not start with advanced machine learning if monitoring data are incomplete or inconsistent. Stage 2: AI-assisted detection introduces dynamic baselines and anomaly scoring but keeps all actions manual. Success is measured by alert quality and whether engineers find model explanations useful.

Stage 3: predictive operations adds forecasting, incident-risk scoring, and probable root-cause ranking. At this stage, the system should integrate with ticketing and change management so predictions are connected to operational workflows. Stage 4: guardrailed automation allows low-risk actions to execute automatically and medium-risk actions under explicit policy. Stage 5: policy-governed autonomous operations uses closed-loop remediation for well-understood scenarios with strong evidence, tested rollback, ongoing model validation, and human exception handling. The transition between stages should be based on measured safety and accuracy rather than a fixed timetable.

Several technical risks require attention. The first is false confidence. High model accuracy on historical data does not guarantee safety under new workloads. Concept drift, platform upgrades, new applications, or changes in user behavior can invalidate learned patterns. Continuous monitoring and retraining governance are therefore required (Cao et al., 2025). The second risk is alert automation without diagnosis. Automating responses to weakly understood anomalies may amplify incidents. Root-cause ranking, dependency context, and precondition checks should precede disruptive action.

The third risk is black-box decision making. If DBAs cannot understand why a model recommends failover or plan intervention, they will either reject the system or trust it inappropriately. Explainable database tuning and anomaly-detection literature shows the importance of interpretable reasoning and human-in-the-loop operation (Ouared et al., 2023; Li et al., 2024). The fourth risk is telemetry and model security. Monitoring data, model pipelines, credentials, and automation accounts are privileged assets. Least privilege, encryption, separation of duties, signed model artifacts, secure secrets management, and immutable audit trails should be part of the platform architecture.

The fifth risk is vendor lock-in or fragmented autonomy. Heterogeneous enterprises should avoid isolated AI silos that cannot correlate cross-platform incidents. A practical compromise is to retain vendor-native autonomous capabilities while maintaining a cross-platform AIOps layer for service visibility, governance, and incident correlation.

The DBA role changes rather than disappears. Automating routine collection, correlation, and first-level diagnosis allows DBAs to focus on architecture, performance engineering, data governance, resilience, capacity strategy, and model oversight. Technology investment should therefore be paired with training in time-series analytics, AI governance, cloud architecture, SRE practices, and secure automation.

VIII. LIMITATIONS AND FUTURE RESEARCH

This paper is a structured review and conceptual framework, not a production benchmark. No Saudi government or enterprise database telemetry was available, so the framework's effect on outage frequency, MTTR, or cost cannot yet be quantified. Vendor capabilities also evolve rapidly, and results from one DBMS or workload may not transfer directly to another. In addition, many anomaly-detection studies use public benchmark datasets that do not fully

represent enterprise database workloads, maintenance windows, change events, or business seasonality.

Future research should validate the framework using anonymized multi-platform telemetry from Saudi organizations. A useful study design would compare static monitoring with AI-assisted monitoring across Oracle, PostgreSQL, and SQL Server environments; measure precision, warning lead time, root-cause accuracy, alert reduction, and MTTR; and evaluate performance under planned workload shifts. Further work should investigate privacy-preserving or federated learning across organizations, causal graph models for cross-service diagnosis, LLM-assisted log interpretation with strict data controls, and reinforcement learning for bounded remediation. Research should also examine how Saudi AI governance and cybersecurity requirements can be translated into machine-enforceable controls for autonomous infrastructure.

IX. CONCLUSION

AI can materially improve database operations when it is used to move monitoring from reactive alarms toward prediction, contextual diagnosis, and safe incident prevention. The literature shows strong progress in AI4DB, AIOps, multivariate anomaly detection, explainability, and database-specific root-cause analysis. However, anomaly detection alone is insufficient for autonomous database management. Production value emerges from integrating telemetry, feature engineering, adaptive models, dependency-aware diagnosis, operational policy, rollback, security, and human oversight.

For Saudi Arabia, this direction is closely aligned with Vision 2030, the National Strategy for Data and AI, digital-government development, cloud adoption, cybersecurity, and responsible AI governance. The proposed six-layer Saudi Autonomous Database Operations Framework provides an implementation path from observability to governed automation, while the maturity model encourages organizations to increase autonomy only when evidence supports safety. The recommended objective is not a database environment without administrators; it is an environment in which AI continuously supports administrators, predicts risk earlier, explains likely causes, automates repetitive low-risk work, and prevents well-understood incidents under controlled policy. Such an approach can strengthen the reliability of the data infrastructure on which Saudi Arabia's increasingly digital economy depends.

REFERENCES

[1] S. Ahmed, M. Singh, B. Doherty, E. Ramlan, K. Harkin, and D. Coyle, "AI for Information Technology Operation (AIOps): A Review of IT Incident Risk Prediction," in *Proc. 2022 9th Int. Conf. Soft Computing & Machine Intelligence (ISCMI)*, pp. 253–257, 2023, doi: 10.1109/ISCMI56532.2022.10068482. Crossref

[2] J. Audibert, P. Michiardi, F. Guyard, S. Marti, and M. A. Zuluaga, "USAD: UnSupervised Anomaly Detection on Multivariate Time Series," in *Proc. 26th ACM SIGKDD Conf. Knowledge Discovery and Data Mining*, pp. 3395–3404, 2020, doi: 10.1145/3394486.3403392. ACM

[3] J. Bogatinovski, S. Nedelkoski, A. Acker, F. Schmidt, T. Wittkopp, S. Becker, J. Cardoso, and O. Kao, "Artificial Intelligence for IT Operations (AIOPS) Workshop White Paper," *arXiv preprint arXiv:2101.06054*, 2021, doi: 10.48550/arXiv.2101.06054. arXiv

[4] Y. Cao, Y. Ma, Y. Zhu, and K. M. Ting, "Revisiting streaming anomaly detection: benchmark and evaluation," *Artificial Intelligence Review*, vol. 58, Art. no. 8, 2025, doi: 10.1007/s10462-024-10995-w. Springer

[5] J. Diaz-De-Arcaya, A. I. Torre-Bastida, G. Zárate, R. Miñón, and A. Almeida, "A Joint Study of the Challenges, Opportunities, and Roadmap of MLOps and AIOps: A Systematic Survey," *ACM Computing Surveys*, vol. 56, no. 4, Art. no. 84, pp. 1–30, 2024, doi: 10.1145/3625289. ACM

[6] Digital Government Authority, *Effective Integration of Cloud Computing in Digital Transformation of the Education Sector*, Version 1.0. Riyadh, Saudi Arabia: Digital Government Authority, 2025. Digital Government Authority Accessed: Sep. 3, 2026.

[7] E. Dritsas and M. Trigka, "Machine learning in modern database systems: Techniques, architectures, and deployment challenges,"

Engineering Applications of Artificial Intelligence, vol. 170, Art. no. 114225, 2026, doi: 10.1016/j.engappai.2026.114225. ScienceDirect

[8] Gao, P. Yuan, S. Han, Y. Liu, X. Xia, H. Zhang, J. Cui, H. Li, and K. Zhao, "Insights into KPI-based performance anomaly detection in database systems: A comprehensive study," *Expert Systems with Applications*, vol. 285, Art. no. 127959, 2025, doi: 10.1016/j.eswa.2025.127959. ScienceDirect

[9] H. Guo, S. Yuan, and X. Wu, "LogBERT: Log Anomaly Detection via BERT," in *Proc. 2021 Int. Joint Conf. Neural Networks (IJCNN)*, pp. 1–8, 2021, doi: 10.1109/IJCNN52387.2021.9534113. Crossref

[10] A. Jamarani, S. Haddadi, R. Sarvizadeh, M. Haghi Kashani, M. Akbari, and S. Moradi, "Big data and predictive analytics: A systematic review of applications," *Artificial Intelligence Review*, vol. 57, Art. no. 176, 2024, doi: 10.1007/s10462-024-10811-5. Springer

[11] Z. Li, Y. Zhu, and M. van Leeuwen, "A Survey on Explainable Anomaly Detection," *ACM Transactions on Knowledge Discovery from Data*, vol. 18, no. 1, Art. no. 23, pp. 1–54, 2024, doi: 10.1145/3609333. ACM

[12] Liu, Z. Yin, C. Zhao, C. Ge, L. Chen, Y. Gao, D. Li, Z. Wang, G. Liang, J. Tan, and F. Li, "PinSQL: Pinpoint Root Cause SQLs to Resolve Performance Issues in Cloud Databases," in *Proc. 2022 IEEE 38th Int. Conf. Data Engineering (ICDE)*, pp. 2549–2561, 2022, doi: 10.1109/ICDE53745.2022.00236. IEEE

[13] B. Milicevic and Z. Babovic, "A systematic review of deep learning applications in database query execution," *Journal of Big Data*, vol. 11, Art. no. 173, 2024, doi: 10.1186/s40537-024-01025-1. Springer

[14] National Transformation Program, *Annual Report 2025*. Riyadh, Saudi Arabia: Saudi Vision 2030, 2026. Saudi Vision 2030 Accessed: Sep. 3, 2026.

[15] P. Notaro, J. Cardoso, and M. Gerndt, "A Survey of AIOps Methods for Failure Management," *ACM Transactions on Intelligent Systems and Technology*, vol. 12, no. 6, Art. no. 81, 2021, doi: 10.1145/3483424. ACM

[16] Oracle, *Autonomous AI Database Documentation*. Oracle Cloud Infrastructure, 2026. Oracle Accessed: Sep. 3, 2026.

[17] Oracle, *Oracle Autonomous Health Framework User's Guide, Release 26.3*. Redwood Shores, CA, USA: Oracle, 2026. Oracle Accessed: Sep. 3, 2026.

[18] A. Ouared, M. Amrani, and P.-Y. Schobbens, "Explainable AI for DBA: Bridging the DBA's experience and machine learning in tuning database systems," *Concurrency and Computation: Practice and Experience*, vol. 35, no. 21, Art. no. e7698, 2023, doi: 10.1002/cpe.7698. Wiley

[19] G. Pang, C. Shen, L. Cao, and A. van den Hengel, "Deep Learning for Anomaly Detection: A Review," *ACM Computing Surveys*, vol. 54, no. 2, Art. no. 38, pp. 1–38, 2021, doi: 10.1145/3439950. ACM

[20] V. G. S. Ruffo, D. M. B. Lent, M. Komarchesqui, V. F. Schiavon, M. V. O. de Assis, L. F. Carvalho, and M. L. Proença Jr., "Anomaly and intrusion detection using deep learning for software-defined networks: A survey," *Expert Systems with Applications*, vol. 256, Art. no. 124982, 2024, doi: 10.1016/j.eswa.2024.124982. ScienceDirect

[21] Saudi Data & AI Authority (SDAIA), *AI Ethics Principles*, Version 1.0. Riyadh, Saudi Arabia: SDAIA, 2023. SDAIA Accessed: Sep. 3, 2026.

[22] Saudi Data & AI Authority (SDAIA), *AI Adoption Framework*. Riyadh, Saudi Arabia: SDAIA, Sep. 2024. SDAIA Accessed: Sep. 3, 2026.

[23] Saudi Data & AI Authority (SDAIA), *National AI Risk Management Framework*. Riyadh, Saudi Arabia: SDAIA, Apr. 1, 2026. SDAIA Accessed: Sep. 3, 2026.

[24] Saudi Data & AI Authority (SDAIA), *National Strategy for Data & AI*. SDAIA Accessed: Sep. 3, 2026.

[25] Saudi Data & AI Authority (SDAIA), "Strategic Objectives." SDAIA Accessed: Sep. 3, 2026.

[26] Saudi Vision 2030, *Vision 2030 Annual Report 2025*. Riyadh, Saudi Arabia: Kingdom of Saudi

Arabia, 2026. Saudi Vision 2030 Accessed: Sep. 3, 2026.

[27] S. Schmidl, P. Wenig, and T. Papenbrock, "Anomaly Detection in Time Series: A Comprehensive Evaluation," *Proceedings of the VLDB Endowment*, vol. 15, no. 9, pp. 1779–1797, 2022, doi: 10.14778/3538598.3538602. VLDB

[28] P. Wenig, S. Schmidl, and T. Papenbrock, "TimeEval: A Benchmarking Toolkit for Time Series Anomaly Detection Algorithms," *Proceedings of the VLDB Endowment*, vol. 15, no. 12, pp. 3678–3681, 2022, doi: 10.14778/3554821.3554873. VLDB

[29] X. Zhou, C. Chai, G. Li, and J. Sun, "Database Meets Artificial Intelligence: A Survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 3, pp. 1096–1116, 2022, doi: 10.1109/TKDE.2020.2994641. Crossref

[30] B. Zou, J. You, Q. Wang, X. Wen, and L. Jia, "Survey on Learnable Databases: A Machine Learning Perspective," *Big Data Research*, vol. 27, Art. no. 100304, 2022, doi: 10.1016/j.bdr.2021.100304. ScienceDirect