Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

AI-Driven Data Quality Monitoring and Anomaly Detection for Financial and Analytics Systems

Mary Sirisha Duggimpudi

Senior Software Engineer, USA

Corresponding Author: Mary Sirisha Duggimpudi**, E-mail:** marysduggimpudi@gmail.com

ABSTRACT

Enterprise financial and analytics platforms depend on continuous, high-fidelity data validation across ingestion, transformation, and reporting layers. Most organizations still keep deterministic reconciliation controls and statistical or machine-learning-based anomaly detection in separate toolchains, an arrangement that leaves each approach blind to the failure class the other is built to catch. This article examines how enterprise data governance programs can structurally combine rule-based validation with adaptive anomaly detection to meet the operational demands of regulated financial and analytics environments. Drawing on established data quality theory, current enterprise governance frameworks, and recent empirical work on machine-learning-based anomaly detection, in which reconstruction-error-based detection methods have reported AUC values as high as 0.963 against conventional sequence-model baselines, the analysis develops a three-layer reference architecture: a deterministic validation layer for schema and completeness checks, a statistical and machine-learning layer for drift and anomaly scoring, and a governance layer that enforces policy and preserves audit accountability. The comparative review finds that layered approaches combining reconciliation-grade controls with adaptive detection consistently report stronger operational outcomes than either method used alone, particularly in reducing undetected data-quality incidents within high-volume transactional environments. The discussion situates this architecture within the practical constraints of enterprise data quality platforms and identifies where purely statistical methods fall short of audit and compliance expectations. The article contributes a practitioner-oriented synthesis intended for data engineering and governance teams responsible for maintaining data integrity in enterprise financial and analytics systems.

KEYWORDS

data quality monitoring, anomaly detection, enterprise data governance, machine learning, financial analytics systems

ARTICLE INFORMATION

ACCEPTED: 04 October 2026 **PUBLISHED:** 04 November 2026 **DOI:** 10.32996/jcsts.2026.8.14

1. Introduction

Enterprise financial and analytics platforms sit under a structural tension that architecture diagrams rarely make explicit: the same pipelines expected to satisfy audit-grade reconciliation requirements must also catch the kind of subtle, statistically defined anomalies that no reconciliation rule can express. Regulated enterprises in banking, telecommunications, and large-scale technology commonly operate across heterogeneous systems—ERP platforms, billing engines, and financial reporting infrastructure—that were never designed to interoperate cleanly [5]. The fragmentation that results produces familiar operational failures: a transaction recorded in one system but unreconciled in another, an asset inventory drifting away from its system of record, an audit-readiness posture that depends on manual reconciliation rather than systematic control [5]. Regulatory inspection data give this problem a concrete scale. A recent PCAOB inspection cycle reported an aggregate Part IA deficiency rate of 39% across inspected firms [5], a figure that reflects reconciliation and control gaps at scale, not a handful of isolated data-entry errors.

A separate and largely disconnected body of work has, in parallel, advanced statistical and machine-learning approaches to anomaly detection in enterprise data streams. Deep autoencoder architectures have proven especially capable here: one recent method targeting enterprise ETL anomaly detection reported reconstruction-error-based detection accuracy exceeding several conventional sequence models across every metric evaluated [10]. Yet this literature typically treats detection performance as a standalone question, evaluated apart from the reconciliation and policy-enforcement obligations that regulated enterprises must also satisfy on a day-to-day basis.

Copyright: © 2026 the Author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) 4.0 license (https://creativecommons.org/licenses/by/4.0/). Published by Al-Kindi Centre for Research and Development, London, United Kingdom.

This article treats that separation as the problem worth solving. The question is not whether rule-based validation or machine-learning-based detection performs better in isolation; that framing has already produced a substantial and largely inconclusive literature. The more useful question is how the two should be structurally combined within a single enterprise data quality architecture, and what governance layer makes the combination auditable rather than merely accurate. The contribution offered here is a three-layer reference architecture, grounded in the literature reviewed in Section 2, that treats deterministic validation, adaptive detection, and governance enforcement as distinct but interdependent layers rather than competing toolchains an organization must choose between.

2. Literature Review

2.1 Foundational Data Quality Theory

Data quality as a formal discipline traces to structural and consumer-oriented framings developed in the mid-1990s, and both strands remain visible in how enterprise data platforms are designed today. Wang and Strong frame data quality along intrinsic, contextual, representational, and accessibility dimensions, arguing that data quality has to be evaluated relative to the task the data support rather than treated as a fixed, absolute property of the data itself [1]. A given customer record, in their framing, can be simultaneously high quality for a billing reconciliation task and low quality for a marketing segmentation task, because the two tasks depend on different dimensions of that same record. A complementary structural approach organizes data quality around the data model, the data values, and the data representation, treating consistency and accuracy as properties that decompose into measurable sub-dimensions rather than single scores [2]. Together, these frameworks established the analytical vocabulary of completeness, accuracy, consistency, timeliness, and validity that governance and monitoring literature still relies on three decades later. Total Data Quality Management extended these dimensional frameworks into a continuous improvement cycle borrowed from total quality management: define, measure, analyze, improve, repeat [3]. That cyclical framing matters for enterprise platforms specifically, because it recasts data quality as an operational discipline requiring ongoing measurement rather than a periodic audit exercise conducted once and filed away.

2.2 Enterprise Data Governance Frameworks

Contemporary enterprise data governance has formalized these dimensional foundations into operational frameworks with defined ownership and accountability. DAMA-DMBOK positions data governance as the central function connecting eleven data management knowledge areas, offering a vendor-neutral vocabulary for governance responsibilities without prescribing any particular tooling or platform [4]. This non-prescriptive quality is deliberate: it lets the framework apply across industries and technology stacks, but it also means the framework alone does not resolve the specific architectural question this article addresses: how deterministic and adaptive methods should be layered together.

More recent architectural work speaks directly to that question. One proposed four-layer governed architecture integrates deterministic reconciliation, statistical anomaly detection, semantic standardization, and security controls into a single methodology explicitly designed to address the kind of audit deficiency findings described above [5]. The architecture is notable for treating reconciliation and anomaly detection as adjacent layers of the same system rather than separate initiatives owned by different teams, a structural choice this article's own three-layer architecture builds on directly. Complementary empirical work models enterprise data asset quality along three interacting dimensions: management capability, standards conformity, and benefit realization. That research finds that management capability exerts the strongest influence on standards conformity, which in turn drives realized business benefit, suggesting that governance maturity is a precondition for, rather than a byproduct of, effective data quality tooling [6]. Within financial services specifically, unified architectures have been proposed that embed rule-based, statistical, and AI-based quality control methods into a single governed layer spanning ingestion, model pipelines, and downstream applications. Deployed in a production financial setting handling streaming and tabular data across multiple asset classes, this kind of unified system reports measurable gains in anomaly detection recall alongside reductions in manual remediation effort [7], a concrete demonstration that the layering this article proposes is not merely theoretically appealing but operationally achievable.

Regulatory pressure has shaped these frameworks directly, and not always in ways the frameworks fully anticipate. Survey-based research on machine learning practitioners operating under EU regulatory requirements finds persistent gaps between established data quality dimensions and the specific documentation expectations of financial and privacy regulation, along with an unmet need for tooling that integrates technical validation with regulatory compliance reasoning rather than treating the two as separate workstreams [8]. That gap between dimensional data quality theory and regulatory operational reality is precisely what motivates the governance layer this article proposes in Section 3.3.

2.3 Machine-Learning-Based Anomaly Detection

A parallel and largely separate literature has advanced statistical and deep-learning methods for anomaly detection in enterprise data streams, generally without engaging the governance concerns discussed above. Data quality dimension taxonomies developed specifically for machine learning contexts categorize quality requirements as intrinsic, contextual, representational, and accessibility-based, echoing the Wang and Strong framework, but applied to ML pipeline inputs rather than to general business data [9]. This convergence is worth noting: the ML-specific literature independently arrived at a dimensional structure the data

quality field had already established decades earlier, which suggests the underlying dimensions are more fundamental than any single application domain.

Deep autoencoder architectures have shown particular strength within this domain. An autoencoder-based method targeting enterprise ETL anomaly detection reported an area under the curve of 0.963 and accuracy of 0.928 on a large-scale dataset drawn from financial transaction and business log records, outperforming 1D-CNN (AUC 0.914), LSTM (AUC 0.926), Transformer (AUC 0.938), and state-space-model (Mamba, AUC 0.945) baselines across all four evaluated metrics [10]. The 1D-CNN baseline itself builds on convolutional anomaly detection work originally developed for manufacturing and equipment-monitoring contexts, where one-dimensional convolutional architectures were shown to detect anomalies effectively in machine-center and CNC-equipment telemetry before the same architectural family was adapted to enterprise data-stream settings [12]. The pattern across these baselines is itself informative: performance improves roughly in step with each architecture's capacity to model long-range and nonlinear dependencies, though the reconstruction-error framing of the autoencoder approach appears to offer an additional advantage specific to anomaly detection tasks, where labeled anomaly examples are typically too scarce to train a purely supervised classifier. Related architectures reinforce this pattern: LSTM-CNN hybrid models applied to streaming anomaly detection [13], stacked transformer representations combined with one-dimensional convolutional networks for time-series anomaly detection [14], and state-space-based unsupervised detection frameworks for multi-class scenarios [15] collectively demonstrate that deep sequence and reconstruction-based models consistently outperform threshold-based statistical baselines in high-volume, high-frequency data environments. Benchmark evaluations of streaming anomaly detection methods further confirm that model performance varies substantially with data non-stationarity [16], a finding with direct operational implications: a detection model tuned and validated against last quarter's transaction patterns may already be degrading against this quarter's data before anyone notices, which is itself the kind of drift a purely static rule-based system cannot catch either.

Despite this technical maturity, practitioner-facing research complicates any assumption that better models alone solve the enterprise problem. Practitioners working on regulatory-aligned data quality in machine learning systems describe their work, in their own words, as largely ad hoc "detective work" rather than a systematized discipline, with limited tooling connecting statistical anomaly detection outputs to the audit and governance documentation regulated enterprises are required to produce [11]. That description is a useful corrective to a purely technical reading of the anomaly-detection literature: the bottleneck is frequently not detection accuracy but the absence of a structured path from a detected anomaly to a documented, auditable resolution. Earlier research into production machine learning systems reached a related conclusion from the infrastructure side, finding that data management rather than model architecture is typically the leading source of failure in deployed ML systems, and that weak data validation allows failures to accumulate silently before surfacing as model degradation [18]. Closing that gap is the specific problem the architecture proposed in Section 3 is built to address. Related reliability engineering practice reinforces the same pattern from a different vantage point: recent work analyzing AI-workload site reliability engineering identifies infrastructure-centric observability—monitoring conventional system metrics while remaining blind to model-specific and data-quality degradation—as a recurring and costly anti-pattern in production AI environments, one that a governance layer bridging statistical detection and operational accountability is specifically positioned to correct [17].

3. Methodology

This article develops its reference architecture through a synthesis method rather than a novel algorithmic contribution: the literature reviewed above, spanning data quality theory, enterprise governance frameworks, and machine-learning anomaly detection techniques, is examined for points of structural compatibility and incompatibility, and the resulting architecture is evaluated against the operational requirements documented directly in the governance and practitioner literature discussed in Section 2. This synthesis approach mirrors the methodology used in comparable enterprise architecture work, where the contribution is an integrative framework rather than a new detection algorithm [5]. It is the appropriate method here because the central problem this article addresses is architectural integration, not algorithmic performance; the algorithms already exist and already perform well individually, as Section 2.3 documents; what is missing is a structural account of how they fit together with governance controls. The three-layer model proposed here is deliberately narrower in scope than the four-layer Governed Enterprise Reconciliation Architecture (GERA) described in [5], which additionally treats semantic standardization and security controls as distinct architectural layers. This article folds those concerns into the governance layer rather than separating them further, and instead concentrates its contribution on the specific interface between deterministic reconciliation and statistical/ML-based anomaly detection, a boundary GERA addresses at the architectural level but does not develop through a worked operational scenario. Section 3.4 below supplies that scenario, illustrating concretely how a detected statistical anomaly moves from the detection layer to a governance-owned remediation outcome, which is this article's specific point of departure from [5] rather than a restatement of it. The proposed architecture separates enterprise data quality management into three interdependent layers, each addressing a distinct failure mode that the others cannot.

3.1 Deterministic Validation Layer

The deterministic layer performs schema validation, completeness checks, and cross-system reconciliation using rule-based logic, including schema validation, null and referential-integrity constraints, duplicate-record rules, and fixed freshness SLAs, precisely the class of controls that regulatory inspection frameworks evaluate directly [5]. This layer operates across tiers of increasing scope.

Row-level validation confirms individual record structure and required-field completeness at the point of ingestion. Aggregation-level validation confirms that summary statistics reconcile across source and target systems once data have moved through transformation. Cross-system reconciliation validation, the broadest tier, confirms that the same underlying business event is represented consistently across otherwise disconnected platforms, the specific failure mode that produces the audit findings described in Section 1. This tiered structure mirrors the reconciliation patterns documented in governed enterprise architecture work, where deterministic checks form the audit-facing backbone that any data quality program must maintain regardless of what adaptive detection capability sits alongside it [5].

3.2 Statistical and Machine-Learning Detection Layer

The second layer applies statistical and machine-learning techniques to catch the anomaly class that deterministic rules cannot express by construction: subtle distributional drift manifesting as unusual transaction volume, distribution shift, seasonality deviation, or changes in null-rate and freshness behavior; an emerging fraud pattern that does not yet violate any individual validation rule; or gradual data degradation that stays within the bounds of every existing check while still representing a meaningful departure from normal system behavior. Reconstruction-error-based deep autoencoder methods suit this layer particularly well, given their demonstrated capacity to model normal data distributions in an unsupervised setting and flag deviations through reconstruction error rather than requiring the labeled anomaly examples that are typically scarce in enterprise data quality contexts [10]. The layer is deliberately positioned downstream of, rather than as a replacement for, the deterministic layer: statistical detection operates on data that have already passed schema and completeness checks, which means its detection surface is confined to the anomaly types those checks cannot see. This positioning also has a practical performance benefit: because the deterministic layer has already filtered out structural failures, missing fields, malformed records, and broken referential integrity, the statistical layer's input stream is cleaner than raw ingestion data would be, which reduces the false-positive burden that unfiltered noise would otherwise place on a reconstruction-error-based detector. A detection model trained on data still containing structural defects risks learning those defects as part of its notion of "normal," which weakens its ability to flag genuine drift later. Sequencing the layers in this order—deterministic first, adaptive second—is therefore not merely an organizational convenience but a modeling requirement.

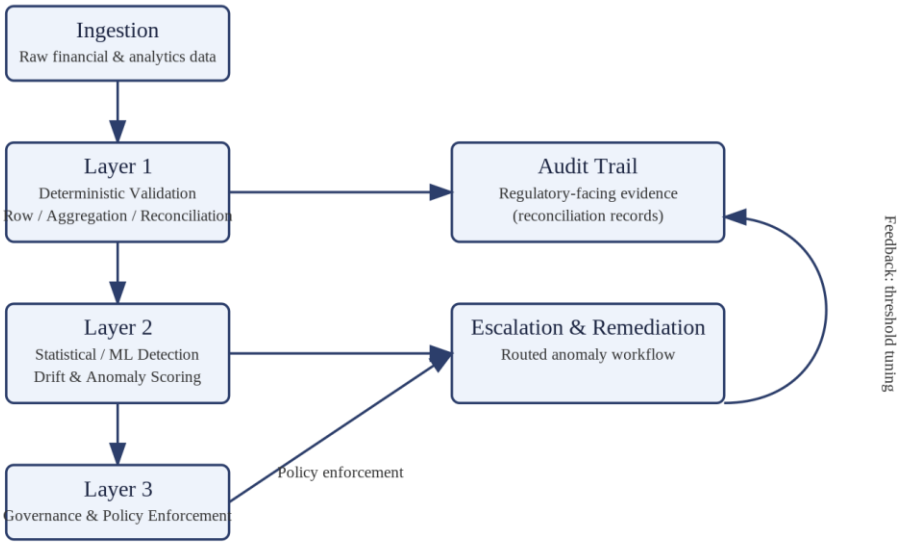


Figure 1. Three-layer enterprise data quality architecture: deterministic validation, statistical/ML detection, and governance/policy enforcement, with feedback-driven threshold tuning.

3.3 Governance and Policy Enforcement Layer

The third layer enforces policy: it sets escalation thresholds, routes detected anomalies to the appropriate remediation workflow, and maintains the audit trail that regulatory review requires. This layer is what closes the gap practitioners describe when they characterize regulatory-aligned quality work as undocumented detective work [11]. Without it, statistical detection produces alerts without accountability; a flagged anomaly with no defined owner or resolution path is operationally indistinguishable from noise, and deterministic validation produces audit trails without adaptive coverage, satisfying an auditor's checklist while missing the drift patterns adaptive detection exists to catch. Unified frameworks that embed all three concerns into a single governed layer report measurable reductions in manual remediation effort specifically because the governance layer routes detected issues automatically rather than requiring manual triage at each step [7], which is the clearest available evidence that this third layer is not organizational overhead but the mechanism that makes the first two layers operationally useful together.

3.4 Illustrative Application to a Cross-System Reconciliation Scenario

A brief scenario illustrates how the three layers interact in a setting typical of the enterprises this article addresses. Consider a financial reporting pipeline that ingests transaction records from a billing system and reconciles them against a downstream general-ledger platform, a pattern consistent with the ERP-to-financial-reporting fragmentation documented in cross-system reconciliation research [5]. At the deterministic layer, row-level validation confirms that each incoming transaction record contains the required fields (transaction identifier, amount, timestamp, and account code) before aggregation-level checks confirm that the day's summarized transaction totals reconcile against the general ledger's corresponding entries. A transaction present in the billing system but absent from the ledger, or vice versa, is flagged immediately as a deterministic reconciliation failure, precisely the class of discrepancy that produces audit findings when left unresolved [5].

Not every failure mode surfaces this cleanly. Suppose transaction volumes from a particular region begin drifting upward over several weeks without violating any individual reconciliation rule; each day's totals still balance, each record still contains its required fields, but the underlying distribution of transaction amounts has shifted in a way that later proves to reflect a misconfigured currency-conversion routine rather than genuine business growth. This is precisely the anomaly class the statistical detection layer exists to catch: a reconstruction-error-based model trained on the historical distribution of transaction amounts would register elevated reconstruction error as the drifted values diverge from previously learned normal patterns, well before the cumulative misstatement became large enough to trigger a downstream reconciliation failure or an external audit finding [10]. The detection layer's output in this scenario is not a pass/fail signal but a graded anomaly score, which is where the governance layer's role becomes decisive: the score must be routed to a data engineering or finance-operations owner, and evaluated against a defined escalation threshold. If confirmed as a genuine issue rather than a false positive, it must be resolved through a documented remediation workflow that itself becomes part of the audit trail [7]. Absent this third layer, the anomaly score alone would have no defined path to resolution, reproducing exactly the unowned-alert problem discussed in Section 3.3.

3.5 False-Positive and Model-Drift Management

Sustaining the statistical detection layer in production requires operational controls beyond the detection algorithm itself. Escalation thresholds calibrated against a single historical baseline degrade as seasonality, holidays, promotional cycles, and other recurring business events shift the definition of "normal" over time; a threshold that correctly separates signal from noise in one reporting period can generate excessive false positives, or miss genuine drift, in the next [21]. Because reconstruction-error-based detectors of the kind described in Section 3.2 are trained on a historical distribution, distinguishing a benign regime shift from a genuine data-quality anomaly requires monitoring mechanisms that are explicitly separable from the anomaly signal itself, rather than folding both into a single static cutoff [22]. Feedback from confirmed and resolved incidents, routed through the governance layer described in Section 3.3, provides the mechanism for recalibration: each remediation outcome—whether it confirms a genuine anomaly or reclassifies an alert as a false positive tied to a known seasonal or business event—should update the detection model's baseline rather than leaving the original threshold in place indefinitely. Absent this feedback loop, a detector tuned against last quarter's transaction patterns is liable to degrade against the current quarter's data well before that degradation becomes visible in aggregate accuracy metrics, compounding the non-stationarity concern already noted in Section 2.3 [16].

4. Results/Findings

Table 1. Comparative synthesis of rule-based, ML-based, and combined data quality approaches.

Approach	Detection Coverage	Audit Traceability	Reported Accuracy/AUC Range
Rule-based validation (deterministic)	Low; cannot express statistical drift	High; directly auditable	N/A (binary pass/fail rules) [5]
ML-based anomaly detection alone	High; captures subtle/multivariate drift	Low; requires translation to audit format	AUC 0.914-0.963 [10],[13]-[16]
Combined / layered (proposed architecture)	High; governance layer routes ML detections	High; deterministic layer preserves audit trail	Reported recall gains + reduced remediation effort [7]

Table 1 synthesizes reported performance characteristics across the three architectural approaches reviewed in Section 2: rule-based validation alone, machine-learning-based anomaly detection alone, and combined or layered approaches, drawn from references [5], [7], [10], [13]–[16]. The pattern across these sources is consistent even though the underlying studies differ in domain and methodology: rule-based validation alone offers strong audit traceability but limited coverage of statistical drift; machine-learning detection alone offers strong drift coverage but weak native audit traceability; combined approaches report the strongest outcomes on both dimensions simultaneously, which is the core empirical claim this article's architecture rests on.

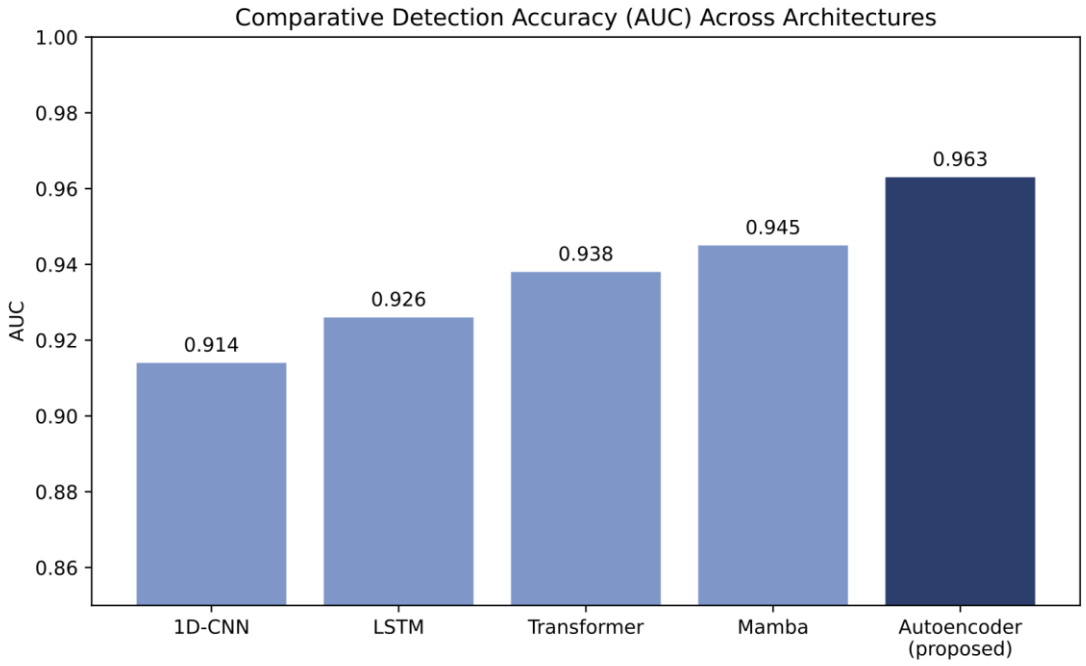


Figure 2. Comparative detection accuracy (AUC) across five architectures evaluated on enterprise ETL anomaly detection (source: reference [10]).

Figure 2 visualizes comparative detection accuracy (AUC) across the deep-learning architectures reviewed in Section 2.3: 1D-CNN, LSTM, Transformer, Mamba, and the autoencoder-based method using the values reported in [10]: 0.914, 0.926, 0.938, 0.945, and 0.963, respectively. The near-monotonic improvement across architectures of increasing representational capacity is itself a finding worth noting, though the practical significance of the gap between, say, 0.945 and 0.963 depends heavily on the cost asymmetry between false positives and false negatives in a given enterprise's specific operational context. Calibrating for this consideration is the responsibility of the governance layer proposed in Section 3.3 rather than the detection layer itself.

Across the reviewed empirical literature, layered or hybrid approaches consistently report stronger outcomes than single-method approaches evaluated on their own. The unified DataOps framework deployed in a production financial environment reported measurable gains in anomaly detection recall alongside reductions in manual remediation effort when rule-based, statistical, and AI-based methods were embedded into a single governed layer rather than operated as separate systems [7]. This pattern is consistent with the governed reconciliation architecture literature, which finds that deterministic reconciliation alone, absent adaptive detection, leaves organizations exposed to precisely the kind of undetected drift that produces audit deficiency findings [5], while adaptive detection alone, absent governance enforcement, produces detected anomalies without a defined remediation or accountability path [11].

Table 2. Data quality dimension mapped to architectural layer and detection mechanism.

Data Quality Dimension	Architectural Layer	Detection Mechanism
Completeness	Layer 1 (Deterministic)	Row/field-level completeness checks [1],[2]
Accuracy	Layer 1 + Layer 2	Reconciliation + reconstruction-error scoring [1],[2],[10]
Consistency	Layer 1 (Deterministic)	Cross-system reconciliation [2],[5]
Timeliness	Layer 2 (Statistical/ML)	Drift and staleness detection [9]
Validity	Layer 1 (Deterministic)	Schema validation [1],[9]
Auditability/Traceability	Layer 3 (Governance)	Policy enforcement, escalation routing [5],[7]

Table 2 maps each data quality dimension established in the foundational literature reviewed in Section 2.1 to the architectural layer and detection mechanism best suited to address it, synthesizing the dimension taxonomy from [1], [2], and [9] against the layered architecture developed in Section 3. This mapping is intended as a practical reference for governance teams deciding where a given data quality concern belongs architecturally, rather than as an additional empirical finding. The mapping also clarifies a point easy to lose in the aggregate performance figures above: no single layer addresses every dimension, and dimensions such as accuracy require both layers acting in sequence rather than either layer acting alone. Timeliness, by contrast, sits almost entirely within the statistical detection layer's domain, since deterministic rules can confirm that a field is populated but cannot easily

express a threshold for "how stale is too stale" without the kind of distributional baseline that drift-detection methods are built to maintain.

5. Discussion

The comparative findings suggest that the choice between deterministic and adaptive data quality methods is a false dichotomy in regulated enterprise contexts, and treating it as a real choice, as much of the existing literature implicitly does by evaluating each approach in isolation, obscures the more useful architectural question. Deterministic validation remains necessary because it produces the auditable, rule-traceable evidence that regulatory review explicitly evaluates [5]; no reconstruction-error score, however accurate, substitutes for a documented reconciliation control when an auditor asks how a specific discrepancy was caught and resolved. At the same time, deterministic rules are structurally incapable of expressing the kind of gradual, multivariate drift that autoencoder-based methods are designed to detect [10], which means an enterprise relying solely on rule-based validation will systematically miss precisely the anomaly class adaptive detection exists to catch.

The practitioner literature suggests this integration gap is not primarily a technical limitation but an organizational and tooling one. Practitioners operating under EU regulatory requirements describe persistent difficulty translating statistical data quality findings into the documentation formats regulatory review expects [8], and separate practitioner research describes regulatory-aligned quality work as ad hoc detective work performed without integrated tooling [11]. Both observations point to the same underlying issue: the technical capability to detect anomalies has outpaced the organizational and architectural capability to make those detections legible to audit and compliance functions. This is consistent with the architecture proposed here. The governance layer is not an afterthought bolted onto a detection pipeline after the fact; it is the layer that makes adaptive detection legible to audit and compliance in the first place, and its absence, rather than any deficiency in detection accuracy, appears to be the more common failure mode in practice.

A second implication concerns organizational ownership rather than architecture in the purely technical sense. The empirical work on enterprise data asset quality found that management capability drives standards conformity, which in turn drives realized benefit [6], a sequence suggesting the three-layer architecture proposed here will underperform if treated as a purely technical deployment rather than as a governance initiative with defined ownership. A statistical detection layer installed without a corresponding governance structure to receive its output produces exactly the kind of unowned alert stream that erodes trust in monitoring systems generally; the architectural contribution proposed in Section 3 is only as effective as the organizational commitment to staff and empower its third layer.

A limitation of this synthesis is that the comparative performance figures reviewed in Section 4 are drawn from studies conducted on different datasets, in different domains, and under different evaluation protocols; the reported AUC and accuracy figures are not directly comparable across studies in a strict statistical sense. They are presented here as evidence of a consistent directional pattern—layered and hybrid approaches outperforming single-method baselines—rather than as a precise quantitative comparison across a controlled experiment. Future empirical work applying a consistent evaluation protocol across the three architectural layers proposed here, within a single regulated enterprise dataset, would strengthen the quantitative basis for the architecture's claimed benefits, and would allow the false-positive and remediation-cost tradeoffs discussed in Section 4 to be quantified directly rather than inferred from separate studies. Earlier work on machine learning pipeline validation similarly found that systematic, end-to-end validation processes catch a substantial share of production failures before deployment when embedded consistently across ingestion and transformation stages, a finding that directly supports the deterministic layer's role in this article's proposed architecture rather than treating it as a legacy control superseded by adaptive detection [19]. More broadly, categorizations of the challenges involved in monitoring deployed machine learning models point to the same multi-layered structure argued for here, distinguishing data-quality degradation, model-performance degradation, and resource-optimization concerns as related but analytically separable monitoring problems that a single detection mechanism is unlikely to address adequately on its own [20].

6. Conclusion

Enterprise financial and analytics data quality is not well served by treating deterministic reconciliation and adaptive anomaly detection as competing approaches, nor by adopting one at the expense of the other. The literature reviewed here indicates that the strongest reported operational outcomes emerge when both are embedded within a single governed architecture: deterministic validation providing the audit-traceable backbone that regulatory review requires, adaptive detection providing coverage against the drift and anomaly classes that rules cannot express by construction, and a governance layer translating detected issues into accountable, documented remediation. For enterprise data governance and engineering teams operating financial or analytics platforms under regulatory scrutiny, this suggests that architectural integration, backed by clear governance ownership, is the more durable path toward measurable reductions in undetected data-quality incidents than continued investment in either method alone. For financial and analytics systems specifically, this hybrid approach improves trust by catching both explicit rule violations and the subtler distributional deviations that technically valid records may still contain.

Statements and Declarations

Author Contributions: The author is solely responsible for conceptualization, literature review, methodology, and the writing of this manuscript.

Acknowledgments: The author used Claude (Anthropic, Claude Sonnet model) to assist with manuscript drafting and language refinement. The author takes full responsibility for the accuracy, originality, and integrity of all content presented in this manuscript, and confirms that all data, analysis, and conclusions reflect independent work and verified sources.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers.

References

[1] R. Y. Wang and D. M. Strong, "Beyond Accuracy: What Data Quality Means to Data Consumers," Journal of Management Information Systems, vol. 12, no. 4, pp. 5-33, 2015. Available: https://www.tandfonline.com/doi/abs/10.1080/07421222.1996.11518099

[2] DAMA Netherlands, "Dimensions of Data Quality (DDQ)," version 1.2, 2020. Available: https://dama-nl.org/wp-content/uploads/2020/09/DDQ-Dimensions-of-Data-Quality-Research-Paper-version-1.2-d.d.-3-Sept-2020.pdf

[3] Richard Y et al., "*A Framework for Analysis of Data Quality Research,*" VOL. 7, NO, 4, AUGUST 1995. Available: https://web.mit.edu/tdqm/www/tdqmpub/SURVEYIEEEKDEAug95.pdf

[4] DAMA International, "DAMA-DMBOK: Data Management Body of Knowledge." Available: https://dama.org/learning-resources/dama-data-management-body-of-knowledge-dmbok/

[5] Z. Qiu, "*Data Engineering Patterns for Cross-System Reconciliation in Regulated Enterprises: Architecture, Anomaly Detection, and Governance,*" arXiv:2604.15108 [cs.DB], 2026. Available: https://arxiv.org/abs/2604.15108

[6] Y. Zhu, A. Wu, R. Zhang, X. Wang, J. Yu, and A. Han, "*Enterprise Data Asset Quality: A Management-Standard Conformity-Benefit Realization Framework and Formation Mechanisms,*" arXiv preprint. Available: https://arxiv.org/pdf/2606.26186

[7] D. Saini, B. Jain, N. Ujjwal, P. Sommer, D. R. Mbanga, and D. Mehta, "A Unified AI System for Data Quality Control and DataOps Management in Regulated Environments," arXiv preprint, 2025. Available: https://arxiv.org/abs/2512.05559

[8] Y. Wang, K. Irion, P. Groth, and H. Harmouch, "Machine Learning Practitioners' Views on Data Quality in Light of EU Regulatory Requirements: A European Online Survey," arXiv preprint, 2026. Available: https://arxiv.org/abs/2602.06594

[9] Y. Zhou et al., "A Survey on Data Quality Dimensions and Tools for Machine Learning," arXiv preprint, 2024. Available: https://arxiv.org/pdf/2406.19614

[10] X. Chen, S. U. Gadgil, K. Gao, Y. Hu, and C. Nie, "Deep Learning Approach to Anomaly Detection in Enterprise ETL Processes with Autoencoders," arXiv preprint, 2025. Available: https://arxiv.org/pdf/2511.00462

[11] Y. Wang et al., "'Detective Work We Shouldn't Have to Do': Practitioner Challenges in Regulatory-Aligned Data Quality in Machine Learning Systems," FAccT '26: The 2026 ACM Conference on Fairness, Accountability, and Transparency, 3217 - 3241, 2026. Available: https://dl.acm.org/doi/10.1145/3805689.3812313

[12] A. Athar, M. A. I. Mozumder, S. Ali, et al., "Deep learning-based anomaly detection using one-dimensional convolutional neural networks (1D CNN) in machine centers (MCT) and computer numerical control (CNC) machines," PeerJ Computer Science, vol. 10, e2389, 2024. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11623112/

[13] A. Duraj, P. S. Szczepaniak, and A. Sadok, "Detection of anomalies in data streams using the LSTM-CNN model," Sensors, vol. 25, no. 5, 1610, 2025. Available: https://www.mdpi.com/1424-8220/25/5/1610

[14] J. Kim, H. Kang, and P. Kang, "Time-series anomaly detection with stacked transformer representations and 1D convolutional network," ACM, vol. 120, 105964, 2023. Available: https://dl.acm.org/doi/10.1016/j.engappai.2023.105964

[15] H. He, Y. Bai, J. Zhang, et al., "*MambaAD: Exploring state space models for multi-class unsupervised anomaly detection,*" arXiv:2404.06564 [cs.CV], 2025. Available: https://arxiv.org/abs/2404.06564

[16] Y. Cao, Y. Ma, Y. Zhu, et al., "Revisiting streaming anomaly detection: benchmark and evaluation," Artificial Intelligence Review, vol. 58, no. 1, 8, 2024. Available: https://link.springer.com/article/10.1007/s10462-024-10995-w

[17] S. Kaimal, "*Avoiding SRE Anti-Patterns in AI Workloads: A Framework for Production-Ready Machine Learning Systems,*" Journal of Computer Science and Technology Studies, vol. 7, no. 10, pp. 504-510, 2025. Available: https://al-kindipublisher.com/index.php/jcsts/article/download/11207/9965/30587

[18] N. Polyzotis, S. Roy, S. E. Whang, and M. Zinkevich, "Data Management Challenges in Production Machine Learning," Google Research. Available: https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/46178.pdf

[19] R. Vadavalasa, "Data Validation Process in Machine Learning Pipeline," ResearchGate, 2021. Available: https://www.researchgate.net/publication/351022721_Data_Validation_Process_in_Machine_Learning_Pipeline

[20] T. Schröder and M. Schulz, "Monitoring machine learning models: a categorization of challenges and methods," Data Science and Management, Volume 5, Issue 3, Pages 105-116, 2022. Available: https://www.sciencedirect.com/science/article/pii/S2666764922000303

[21] M. A. Li and A. Gautam, "*Segmented Confidence Sequences and Multi-Scale Adaptive Confidence Segments for Anomaly Detection in Nonstationary Time Series,*" arXiv:2508.06638 [cs.LG], 2025. Available: https://arxiv.org/abs/2508.06638

[22] J. Park, F. Chiang, and M. Milani, "*Adaptive Anomaly Detection in the Presence of Concept Drift: Extended Report,*" arXiv:2506.15831 [cs.DB], 2025. Available: https://arxiv.org/abs/2506.15831