



International Journal of Research Publication and Reviews

Journal Homepage: www.ijrpr.com ISSN 2582-7421

An Autonomous Control Framework for Data Quality Governance in Trading and Market Risk Monitoring

*Anubhav Birendra Jha**

Pune Maharashtra India

DOI : <https://doi.org/10.55248/gengpi.07.1026.3234>

ABSTRACT

Financial institutions increasingly depend on large, fast-moving volumes of reference, pricing, transaction, and valuation data to calculate market risk exposures, including Value-at-Risk (VaR), sensitivities, stress measures, and regulatory capital. Conventional data quality practice is often organized around reactive detection: controls identify exceptions after data have entered a processing chain, while remediation depends on manual investigation. This paper develops the Autonomous Control Framework (ACF), a conceptual maturity model describing progression from manual detection to rule-based monitoring, automated alerting, predictive anomaly detection, and bounded autonomous remediation at the level of Critical Data Elements (CDEs). The framework is grounded in BCBS 239 risk data aggregation principles, data quality and governance literature, and model risk management considerations. A key distinction is made between autonomous control execution and autonomous judgment: Stage 5 permits only pre-authorized, reversible, auditable actions such as quarantine or use of a documented fallback source, and does not delegate discretionary risk judgment to a machine. The paper is conceptual and illustrative rather than empirically validated. Its principal contribution is a structured bridge between CDE governance, detection latency, market-risk data propagation, and bounded automation, supported by explicit maturity dimensions, an autonomy-eligibility matrix, and testable research propositions. Future research should validate the framework using institutional data, controlled experiments, and expert-based assessment of maturity and autonomy eligibility.

Keywords: Data Quality Management, Data Governance, Critical Data Elements, Market Risk, BCBS 239, Model Risk; Anomaly Detection, Autonomous Controls, Risk Data Aggregation, Data Lineage.

Introduction

Background

The global financial crisis of 2007-2009 exposed weaknesses in the ability of large banking institutions to aggregate risk exposures accurately and rapidly across business lines and legal entities. In response, supervisors placed risk data aggregation, governance, accuracy, completeness, and timeliness at the center of post-crisis reform. BCBS 239 remains a foundational framework for this agenda. Recent supervisory communication confirms that implementation remains an ongoing challenge, particularly around governance, lineage, fragmented data estates, and the use of emerging technologies.

Market-risk processes are especially sensitive to data quality because pricing, curves, volatilities, reference data, positions, collateral inputs, and other critical inputs feed high-frequency valuation and risk calculations. A data-quality defect may therefore propagate through valuation, sensitivities, VaR, stress calculations, capital measures, and reporting before a conventional exception process has completed investigation.

Problem Statement

The central problem addressed in this paper is the detection-lag problem: the interval between the time a data-quality error enters a risk-data environment and the time a control identifies and contains it. Reactive controls can remain effective for many use cases, but their operating model is structurally limited when the time available for detection and remediation is shorter than the downstream calculation cycle. The research problem is therefore not simply how to detect more errors, but how to determine when automation can move detection and bounded remediation closer to the point of data entry without creating uncontrolled model, operational, or governance risk.

Research Gap

Existing literature provides mature concepts for data quality dimensions, data governance, risk data aggregation, and anomaly detection. However, there remains a narrower practitioner-oriented gap: the explicit connection between CDE-level governance maturity and market-risk model integrity. This paper organizes that gap as a maturity and decision framework rather than treating automation as a binary capability.

Research Questions

RQ1: How can data-quality control maturity for market-risk CDEs be represented as a staged framework?

RQ2: Which governance and control dimensions should determine progression between maturity stages?

RQ3: Under what conditions is autonomous remediation appropriate for a market-risk CDE?

RQ4: Which measurable outcomes can be used to evaluate whether higher maturity reduces detection lag and downstream risk-data impacts?

Contributions

- A five-stage Autonomous Control Framework (ACF) linking detection, response, automation, and governance maturity.
- A multidimensional maturity model covering detection, response, decision automation, lineage, governance, and auditability.
- A bounded autonomy-eligibility matrix that separates autonomous control execution from discretionary human judgment.
- A CDE-level application model covering pricing, yield-curve inputs, instrument/counterparty reference data, and collateral/margin inputs.
- A measurement and validation framework that converts the conceptual model into testable research propositions.

Regulatory Context and Literature Review

BCBS 239 and Risk Data Aggregation

BCBS 239 established principles addressing governance, infrastructure, accuracy, completeness, timeliness, and adaptability of risk data. Basel Committee progress reports have continued to identify substantial implementation work across global systemically important banks, and a January 2026 Basel Committee newsletter further emphasized ongoing challenges involving data lineage, cross-border implementation, fragmented data estates, and emerging technologies, including early-stage use of AI in risk data aggregation. These developments support the relevance of a control framework that treats lineage, ownership, quality evidence, and automation as connected capabilities rather than isolated controls.

Model Risk Management

Model-risk governance is relevant because Stage 4 anomaly-detection components can themselves become models or model-like decision systems. For a 2026 publication, SR 11-7 should be treated as historical context: on April 17, 2026, the U.S. banking agencies (the Federal Reserve, OCC, and FDIC) jointly issued Revised Guidance on Model Risk Management (SR 26-2), which supersedes SR 11-7 and emphasizes a risk-based, proportionate approach to model-risk management tailored to an institution's size, complexity, and model risk profile. The revised guidance notably places generative and agentic AI outside its direct scope while stating that existing risk-management principles should still guide governance for such systems. Accordingly, the ACF treats anomaly detection, threshold calibration, validation, monitoring, change control, and override behavior as governance considerations rather than purely technical features.

Data Quality and Data Governance

The data-quality literature distinguishes dimensions such as accuracy, completeness, timeliness, consistency, and fitness for use. Data governance literature additionally emphasizes ownership, stewardship, definitions, lineage, accountability, and control evidence. The ACF translates these general principles into a CDE-level maturity construct in which the ability to automate safely depends on the quality of the governance environment surrounding the data element.

Anomaly Detection and Financial Data

Anomaly-detection research demonstrates the usefulness of statistical and machine-learning approaches for identifying unusual observations in financial systems. Recent literature also identifies limitations around data availability, dimensionality, generalization, and effectiveness measurement. For market-risk data, the practical implication is that Stage 4 should not be defined simply as using AI; it requires stable baselines, monitored performance, explainable alert logic where appropriate, and explicit treatment of regime changes and false positives.

Autonomous and Self-Healing Controls

The term self-healing is used in this paper in a deliberately narrow operational sense. An autonomous control is a pre-authorized mechanism that detects a defined data-quality condition and executes a bounded, reversible, documented corrective action without requiring real-time human approval. This does not mean that the system independently determines what constitutes acceptable market risk or exercises discretionary business judgment. The distinction is central to the governance design of Stage 5.

Research Gap and Conceptual Positioning

The literature and regulatory sources establish the importance of risk-data quality, governance, lineage, and automation, but do not by themselves provide a unified CDE-level maturity model for deciding when predictive detection should progress to bounded autonomous remediation. The ACF is positioned as a conceptual bridge among these streams.

Research Methodology

Research Design

This study adopts a conceptual framework-development design. It does not claim empirical validation of the ACF. The framework is constructed by synthesizing regulatory principles, established data-quality and governance concepts, anomaly-detection literature, and the operational characteristics of market-risk data.

Literature and Regulatory Review

The review draws on core regulatory and academic references and extends them with recent regulatory material and recent financial anomaly-detection research. The review focuses on four evidence streams: (1) risk-data aggregation and reporting; (2) data-quality and governance; (3) model and AI-related risk; and (4) automated/anomaly-based controls in financial systems.

CDE Selection Logic

CDE categories were selected as material inputs to market-risk calculation chains: pricing and market data, yield-curve/rate construction inputs, instrument and counterparty reference data, and collateral/margin valuation inputs. The examples are generalized and illustrative rather than institution-specific.

Framework Development Procedure

- Map the market-risk data flow from source systems to downstream risk and reporting consumers.
- Identify representative CDE failure modes and their potential propagation paths.
- Define five maturity stages according to detection and remediation capability.
- Add cross-cutting maturity dimensions for governance and control evidence.
- Define eligibility conditions for autonomous remediation.
- Specify measurable outcomes and research propositions for later empirical validation.

Validation Strategy

Future empirical validation should combine expert assessment, historical incident data, and controlled back-testing. A suitable validation design would compare CDEs at different maturity levels using detection latency, false-positive rates, remediation time, downstream breaks, and the frequency or magnitude of risk metric restatements. Controlled experiments can additionally test whether approved fallback actions reduce propagation without introducing new material errors.

Market-Risk Data Quality Problem

Representative Data Sources and CDEs

CDE Category	Representative Inputs	Typical Failure Modes	Potential Downstream Effect
--------------	-----------------------	-----------------------	-----------------------------

CDE Category	Representative Inputs	Typical Failure Modes	Potential Downstream Effect
Pricing / market data	Prices, spreads, volatilities, market quotes	Stale, erroneous, unmapped, outlier values	Valuation and sensitivity distortion
Yield curve / rates	Curve nodes, fixings, interpolation inputs	Missing nodes, wrong tenor, broken relationships	Broad repricing and risk-measure distortion
Instrument / counterparty reference	Identifiers, static attributes, hierarchy	Mismapping, unauthorized change, stale attributes	Booking, aggregation, netting or exposure errors
Collateral / margin	Haircuts, collateral values, eligibility inputs	Incorrect valuation, stale values, disputed data	Net exposure and capital impacts

Error Propagation

A useful representation of the propagation chain is: Source, Ingestion, Transformation/Mapping, CDE Quality Gate, Valuation/Risk Engine, Aggregation, and Reporting. A defect can propagate whenever a quality gate is absent, delayed, poorly calibrated, or unable to contain the affected data before downstream use. Figure 1 depicts this chain.

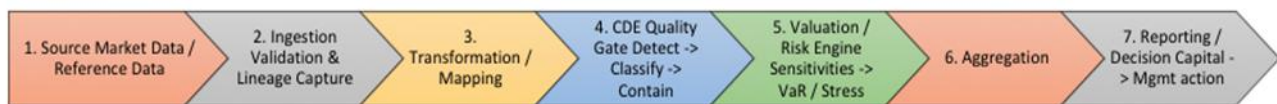


Figure 1. Conceptual market-risk data-quality propagation chain. The figure is conceptual and does not represent a specific institution's architecture.

Detection Lag

Detection lag is defined as the elapsed time between error introduction and control identification. The framework proposes that this lag should be measured directly rather than inferred from the existence of a control. A control can be automated and still have poor practical performance if it runs too late, generates excessive false positives, or lacks authority to contain the affected data.

The Autonomous Control Framework

Five-Stage Maturity Model

The ACF describes five stages of data quality control maturity, distinguished by the degree to which detection and remediation are automated and by whether controls act before or after data enters downstream calculations. Figure 2 shows the progression.



Figure 2. The five-stage ACF maturity progression.

Stage	Name	Core Capability	Remediation	Governance Requirement
1	Manual Detection	Human review / downstream discovery	Manual	Informal or basic ownership
2	Rule-Based Monitoring	Static completeness, range, format and reconciliation rules	Manual	Defined rules and exception ownership

Stage	Name	Core Capability	Remediation	Governance Requirement
3	Automated Detection with Alerting	Real/near-real-time workflow alerts and escalation	Human-led	CDE owner, escalation and SLA
4	Predictive and Anomaly-Based Detection	Dynamic baselines, peer comparison, anomaly scoring	Human-led or controlled workflow	Model validation, monitoring and drift management
5	Autonomous Self-Healing Controls	Pre-authorized detection plus bounded remediation/quarantine	Automated, reversible, auditable	Approved autonomy policy, lineage, logging, rollback and human oversight of rules

Maternity Dimensions

Dimension	Stage 1	Stage 3	Stage 4	Stage 5
Detection	Manual	Automated rules	Predictive / anomaly	Predictive + deterministic trigger
Response	Ad hoc	Workflow escalation	Prioritized response	Pre-authorized response
Decision automation	None	Alert routing	Risk-based prioritization	Bounded action execution
Lineage	Partial / manual	Traceable CDE	End-to-end monitored	End-to-end evidence captured
Governance	Informal	Named owner	Model + data governance	Explicit autonomy governance
Auditability	Case notes	Workflow log	Model + alert evidence	Immutable action and rollback evidence

Governance Gates for Progression

- Stage 2 → 3: named CDE owner, documented rule logic, alert routing, escalation path, and service-level expectations.
- Stage 3 → 4: sufficient historical data, model-development/validation capability, monitoring thresholds, and documented performance metrics.
- Stage 4 → 5: deterministic remediation logic, low and bounded materiality, reversibility, complete lineage, approved fallback, audit logging, rollback, and explicit human escalation.

Autonomous Control Versus Autonomous Judgment

ACF Stage 5 is intentionally narrower than general-purpose autonomous decision-making. The system may execute an approved control action, but it should not independently decide to change a risk methodology, override a valuation judgment, alter a model assumption, or accept an unexplained material exception. Human governance remains responsible for defining and approving the action space.

Autonomy Eligibility Matrix

Not every CDE should be eligible for Stage 5. Eligibility should be assessed before autonomous remediation is enabled and revisited after material changes in data sources, products, models, or market conditions. Figure 3 presents the eligibility logic as a decision sequence.

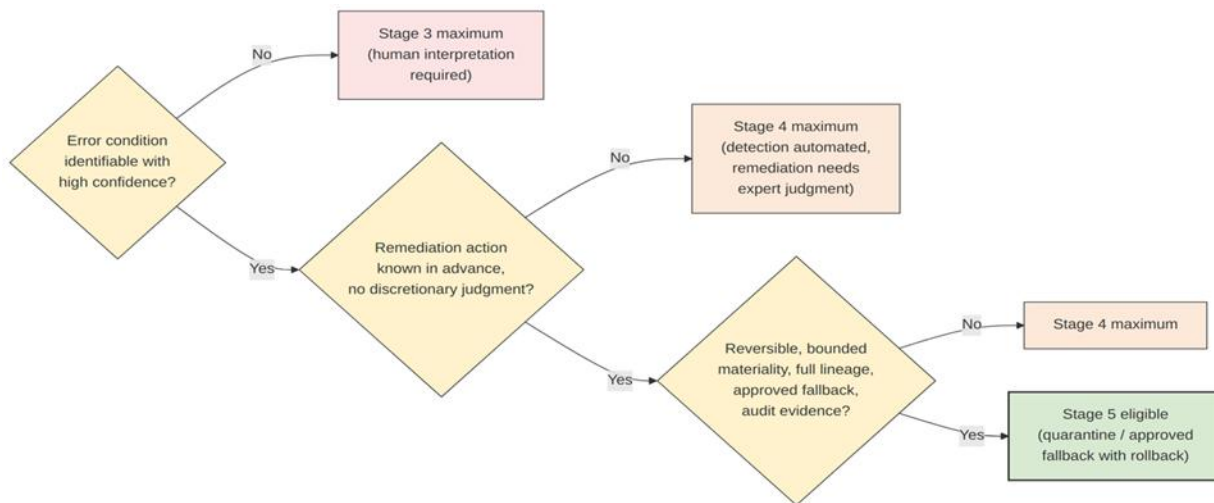


Figure 3. Autonomy eligibility decision tree, translating the eligibility criteria in Section 6 into a sequential decision path.

Criterion	Question	Stage 5 Expectation
Error determinacy	Can the error condition be identified with high confidence?	Yes; condition is explicitly specified
Remediation determinacy	Is the correct action known in advance?	Yes; no discretionary judgment
Reversibility	Can the action be rolled back?	Yes; rollback tested
Materiality	Could the automated action create material risk if wrong?	Low/bounded residual risk
Lineage	Can source, transformation and downstream use be traced?	End-to-end lineage available
Fallback	Is there an approved safe alternative?	Documented and time-bounded
Auditability	Can the system evidence why and how it acted?	Complete action log
Escalation	Can exceptions move immediately to a human owner?	Yes; explicit SLA
Model risk	Does Stage 5 depend on a predictive model?	Autonomy should not rely on ungoverned model judgment

Eligibility Classification

Classification	Meaning	Typical Example
Stage 5 eligible	Deterministic, reversible, low-judgment action is pre-approved	Quarantine anomalous price and temporarily use an

Classification	Meaning	Typical Example
		approved fallback source
Stage 4 maximum	Detection can be automated, but remediation requires expert judgment	Yield-curve construction anomaly requiring curve methodology review
Stage 3 maximum	Even detection/response requires substantial human interpretation or weak data evidence	Complex collateral dispute or ambiguous reference-data correction

Application to Market-Risk CDEs

CDE	Failure Scenario	Detection	Possible Remediation	Human Judgment	Indicative Target
Pricing data	Price deviates materially from trusted sources and expected movement	Rules + peer comparison + anomaly score	Quarantine; approved fallback for bounded period	Confirm source / instrument context	Stage 4; selective Stage 5
Yield curve inputs	Curve node inconsistent with neighboring tenors	Shape checks + cross-curve anomaly detection	Hold / quarantine node; rebuild under approved rule	Review curve construction and market context	Stage 4
Instrument reference	Unauthorized static attribute change	Event-based rules + approval comparison	Revert to last approved value	Investigate authorization and business impact	Stage 3; selective Stage 5
Counterparty reference	Identifier or hierarchy mapping error	Consistency and reconciliation checks	Quarantine mapping; route to steward	Confirm legal/entity mapping	Stage 3-4
Collateral / margin	Valuation or haircut input anomalous	Rules + peer/historical anomaly detection	Quarantine or use approved fallback where permitted	Resolve dispute / eligibility judgment	Stage 4; cautious Stage 5

Illustrative Control Architecture

Figure 4 presents the autonomous control decision architecture. The architecture is intentionally designed so that autonomous action is the result of a governance gate, not merely a model score.

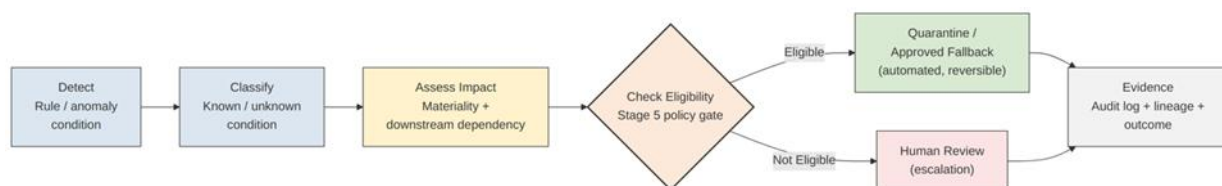


Figure 4. Autonomous control decision architecture: Detect → Classify → Assess Impact → Check Eligibility → Action (human review, or quarantine/approved fallback) → Evidence.

CDE Prioritization Model

Investment prioritization should combine business impact with feasibility of safe automation. A simple conceptual prioritization score can be expressed as: $\text{Priority} = \text{Impact} \times \text{Propagation Exposure} \times \text{Detection-Lag Exposure} \times \text{Automation Feasibility}$. This expression is a research construct, not a validated quantitative scoring model. Institutions should calibrate the components using their own materiality frameworks and control inventories.

Factor	Indicative Measures
Impact	P&L sensitivity, capital relevance, regulatory reporting relevance
Propagation exposure	Number and materiality of downstream consumers
Detection-lag exposure	Time between source arrival, risk calculation and control execution
Automation feasibility	Determinacy, reversibility, lineage, fallback and governance readiness

Measurement and Evaluation Framework

Metric	Definition	Expected Use
Detection latency	Time from error introduction to identification	Assess reduction in detection lag
Containment latency	Time from identification to quarantine/containment	Assess control responsiveness
Precision	Share of alerts that are confirmed issues	Assess alert quality
Recall	Share of known issues detected	Assess coverage
False-positive rate	Non-issues incorrectly flagged	Assess alert fatigue
Mean time to remediation	Average time from detection to correction	Assess operational efficiency
Escalation rate	Share of cases requiring human intervention	Assess autonomy boundary
Downstream P&L breaks	Number/magnitude of breaks attributable to data quality	Assess business effect
VaR / risk restatements	Frequency/magnitude attributable to data defects	Assess risk integrity
Autonomous success rate	Share of automated actions completed without reversal	Assess Stage 5 effectiveness
Rollback / override rate	Share of autonomous actions reversed or overridden	Assess residual risk
Audit completeness	Share of actions with full evidence and lineage	Assess governance quality

Proposed Validation Design

A future empirical study could construct a panel of CDEs across business units and observation periods. Each CDE would receive a maturity-stage assessment using the six ACF dimensions. Outcomes would then be tested using time-based operational measures and downstream risk impacts. Where institutional data are unavailable, a controlled synthetic-data experiment can inject known defects into pricing, curve, and reference-data streams and compare control performance across maturity stages.

Research Propositions

P1: Higher ACF maturity is associated with lower detection latency for market-risk data-quality incidents.

P2: Higher CDE governance maturity is associated with fewer downstream risk-calculation breaks attributable to data-quality defects.

P3: The effectiveness of Stage 4 anomaly detection depends on calibration quality, stability of the underlying data-generating process, and control of false-positive rates.

P4: Stage 5 autonomous remediation is more feasible when error conditions and remediation actions are deterministic, reversible, and low in discretionary judgment.

P5: Stronger data lineage, ownership, and escalation mechanisms increase the feasibility of progression from Stage 3 to Stage 4 and Stage 5.

P6: Excessive false-positive rates reduce operational effectiveness by increasing alert fatigue and weakening human attention to genuine exceptions.

Governance and Implementation

Accountability and Ownership

Each CDE should have an accountable owner, operational steward, control owner, and defined escalation authority. Ownership should cover both the data element and the remediation policy. A central data governance function can coordinate standards, while business and risk functions retain responsibility for the meaning and materiality of the data.

Lineage and Evidence

Autonomous controls require evidence that connects the source record, quality decision, remediation action, downstream consumer, and final outcome. The control record should therefore preserve timestamp, CDE identity, source, rule/model version, decision reason, action, fallback value if any, approver/policy reference, rollback status, and escalation outcome.

Model Risk and Drift

Stage 4 models should be subject to documented development, validation, performance monitoring, change control, and periodic review, consistent with SR 26-2's risk-based and proportionate approach. Drift monitoring should consider changes in market regime, source composition, volatility, product mix, and missing-data patterns. A model should not automatically become an authority merely because it produces a high anomaly score.

Cybersecurity and Change Management

Autonomous remediation expands the control plane and therefore creates additional attack and failure surfaces. Access to remediation policies, fallback sources, model versions, and override mechanisms should be tightly controlled. Changes to autonomy rules should undergo documented testing, approval, deployment controls, and post-change monitoring.

Human Oversight

Human oversight should focus on policy, exceptions, materiality, model governance, and periodic challenge rather than requiring a person to approve every low-risk deterministic action. This preserves operational speed while retaining accountability for the design of the autonomous action space.

Discussion

Theoretical Implications

The ACF reframes data-quality maturity as a multidimensional capability rather than a technology ladder. Automation is treated as the outcome of governance readiness, control evidence, and remediation determinacy. This provides a conceptual mechanism for linking traditional data-quality dimensions with operational risk and model-integrity outcomes.

Practical Implications

For practitioners, the framework provides a roadmap for deciding where automation may be useful and where human judgment should remain central. It also makes the investment case more testable by linking maturity to measurable outcomes such as detection latency, remediation time, and downstream breaks rather than claiming a generalized return on investment without empirical evidence.

Regulatory Implications

The framework is consistent with the continuing emphasis on accurate, timely, governed, and traceable risk data. It should not be interpreted as a regulatory compliance methodology or as evidence that autonomous controls are required. Rather, it provides a conceptual lens through which institutions can document how control automation is bounded, governed, and evidenced.

Limitations and Future Research

- The study is conceptual and has not been empirically validated using proprietary banking data.
- The CDE examples are generalized and may not capture institution-specific product, system, or jurisdictional differences.
- The proposed maturity dimensions and autonomy criteria have not yet been tested for inter-rater reliability or predictive validity.
- The prioritization expression is conceptual and should not be treated as a validated scoring methodology.
- Anomaly-detection effectiveness may deteriorate during structural market regime changes, new-product introductions, or major source-system changes.
- Stage 5 controls may create new operational and cybersecurity risks if fallback data, policy logic, or authorization mechanisms are compromised.

Future research should test the propositions using multi-bank or multi-business-unit evidence where possible; develop validated scoring scales for the six maturity dimensions; examine the relationship between maturity and risk-measure restatements; and investigate how governance arrangements affect the safe deployment of bounded autonomy. Extensions to credit, liquidity, treasury, and regulatory reporting data are also possible.

Conclusion

This paper develops the Autonomous Control Framework as a conceptual maturity model for data-quality controls in trading and market-risk environments. The framework moves from manual detection through rule-based monitoring, automated alerting, predictive anomaly detection, and finally bounded autonomous remediation for a carefully defined subset of CDEs. Its central design principle is that autonomy must be earned through governance: clear ownership, reliable lineage, deterministic remediation, reversibility, materiality controls, auditability, escalation, and model-risk oversight.

The framework therefore does not advocate automation for its own sake. It provides a structured way to identify where earlier detection and bounded automated containment may reduce detection lag and downstream data-quality impacts while preserving human accountability for discretionary risk judgments. The next research step is empirical validation using operational control data and controlled experiments.

Acknowledgment

This work is supported in An Autonomous Control Framework for Data Quality Governance in Trading and Market Risk Monitoring of any state in India. Authors are thankful to References and self-pace learning to carry out the research work.

References

- [1] Basel Committee on Banking Supervision, "Implementation of the Principles for effective risk data aggregation and risk reporting BCBS 239 Principles," Newsletter No. 36, Bank for Int. Settlements, Basel, Switzerland, Jan. 6, 2026
- [2] Board of Governors of the Federal Reserve System, Office of the Comptroller of the Currency, and Federal Deposit Insurance Corporation, "Revised guidance on model risk management," SR 26-2, Washington, DC, USA, Apr. 17, 2026.
- [3] European Central Bank, "Guide on effective risk data aggregation and risk reporting," ECB Banking Supervision, Frankfurt am Main, Germany, May 2024.
- [4] European Parliament and Council of the European Union, "Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)," *Official J. Eur. Union*, vol. L, Jul. 12, 2024.
- [5] Basel Committee on Banking Supervision, "Progress in adopting the Principles for effective risk data aggregation and risk reporting," Bank for Int. Settlements, Basel, Switzerland, Nov. 2023.

-
- [6] S. E. Whang, Y. Roh, H. Song, and J.-G. Lee, "Data collection and quality challenges in deep learning: A data-centric AI perspective," *VLDB J.*, vol. 32, no. 4, pp. 791–813, Jul. 2023.
- [7] National Institute of Standards and Technology, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," NIST AI 100-1, Gaithersburg, MD, USA, Jan. 2023.
- [8] J. Martins, H. S. Mamede, and J. Correia, "Risk compliance and master data management in banking – A novel BCBS 239 compliance action-plan proposal," *Heliyon*, vol. 8, no. 6, Art. no. e09627, Jun. 2022.
- [9] L. Ehrlinger and W. WoB, "A survey of data quality measurement and monitoring tools," *Front. Big Data*, vol. 5, Art. no. 850611, Mar. 2022.
- [10] S. O. Pinto and V. A. Sobreiro, "Literature review: Anomaly detection approaches on digital business financial systems," *Digit. Bus.*, vol. 2, no. 2, Art. no. 100038, 2022.
- [11] G. Pang, C. Shen, L. Cao, and A. van den Hengel, "Deep learning for anomaly detection: A review," *ACM Comput. Surv.*, vol. 54, no. 2, pp. 1–38, Mar. 2021.