International Journal of Multidisciplinary Research and Growth Evaluation.

Auditing Data Governance for AI/ML in Financial Institutions: Verifying the Integrity, Traceability, and Lineage of Training and Production Data under Regulatory Mandates

Puneet Redu
CQF (Certificate in Quantitative Finance), UK
FRM (Financial Risk Manager), USA

* Corresponding Author: **Puneet Redu**

Article Info

ISSN (Online): 2582-7138
Impact Factor (RSIF): 8.04
Volume: 07
Issue: 03
May-June 2026
Received: 17-02-2026
Accepted: 19-03-2026
Published: 21-04-2026
Page No: 42-56

Abstract

The increasing integration of artificial intelligence (AI) and machine learning (ML) into financial institutions has shifted the primary locus of model risk from algorithmic implementation toward the quality, provenance, and governance of data. While traditional model risk management frameworks emphasize conceptual soundness, performance validation, and outcome monitoring, they often treat data governance as a supporting operational function rather than as a core object of independent assurance. This creates a structural gap: institutions may possess formal data governance policies and advanced data architectures, yet lack verifiable mechanisms to demonstrate to auditors and regulators that training and production data are complete, traceable, reproducible, and appropriately controlled throughout the model lifecycle.

This paper proposes a structured, audit-oriented framework for data governance in AI/ML systems within financial institutions. The central conceptual shift is to treat datasets not merely as operational inputs, but as controlled model artifacts subject to explicit versioning, immutability, provenance tracking, and independent verification. By reframing data governance as an assurance and verification problem rather than solely a management or architectural problem, the framework translates high-level regulatory expectations into concrete control objectives, audit tests, and evidence standards.

The paper introduces the Transparent Extract Transform Load (T-ETL) architecture, an extension of conventional data pipelines that embeds lineage capture, policy enforcement, and integrity verification directly into data ingestion, transformation, and deployment processes. T-ETL integrates graph-based lineage representations, cryptographic hash commitments, and bi-temporal data reconstruction to support reproducibility and forensic auditability. This architecture enables auditors and supervisors to reconstruct the precise data state, transformation logic, and governance controls in effect at the time of any model decision.

Mathematically, the framework formalizes data pipelines as transformation functions over versioned datasets, models lineage as directed acyclic graphs over data artifacts, and defines integrity, drift, and bias as computable properties subject to continuous monitoring. These formalizations support the transition from qualitative governance assertions to quantitative, testable assurance mechanisms.

By mapping these technical controls to regulatory mandates including SR 11-7, BCBS 239, the EU Artificial Intelligence Act, and the Digital Operational Resilience Act (DORA), the paper provides a coherent audit framework that aligns operational data practices with supervisory expectations. The contribution is not the introduction of new regulatory principles, but the operationalization of existing ones into a unified, model-centric assurance structure. This approach supports regulatory defensibility, enhances institutional accountability, and strengthens trust in AI-driven financial decision-making.

DOI: https://doi.org/10.54660/.IJMRGE.2026.7.3.42-56

Keywords: Data Governance, AI Auditing, AI Governance, Model Risk Management (MRM), SR 11-7, BCBS 239, EU AI Act, Data Lineage, Transparent ETL, Operational Resilience, Algorithmic Fairness

1. Introduction

1.1. The Changing Nature of Model Risk in Finance

Artificial intelligence and machine learning have transitioned from experimental technologies to core infrastructure within

financial institutions. They are now embedded in credit underwriting, fraud detection, market surveillance, stress testing, anti-money laundering, and customer interaction systems *[1, 2]*. This transformation has altered not only the tools of finance, but the nature of financial risk itself. Traditional quantitative models rely on explicitly specified equations and assumptions. In contrast, modern AI systems derive behavior from data through statistical learning processes, making the characteristics of the data as important as, and often more important than, the model architecture. [3]
Let a model be represented as a function:

$$f\theta: X \rightarrow Y$$

where X is the input feature space, Y is the output space, and θ denotes learned parameters. In classical modeling, risk arises primarily from misspecification of f. In ML-based systems, however, risk arises jointly from fθ and from the empirical distribution P(X) induced by the training and production data *[4]*. Changes to P(X) — whether through data drift, bias, corruption, or undocumented transformation — can materially alter model behavior even if θ remains unchanged. [4]
This dependence creates a structural shift: data becomes part of the model itself. Consequently, the governance of data becomes inseparable from the governance of models. [1] [5]

Because model behavior is inextricably linked to the empirical input distribution P(X), the T-ETL framework proposed in this paper is designed to treat the governed realization of P(X) —as instantiated through versioned datasets and transformations—as a first-class, model-linked artifact subject to explicit control, traceability, and independent verification

1.2. Regulatory Expectations and the Governance Gap

Regulatory frameworks such as SR 11-7, BCBS 239, the EU AI Act, and DORA articulate principles of transparency, accountability, integrity, and resilience. However, these frameworks largely specify what institutions must achieve rather than how they should demonstrate achievement. [1, 2, 5, 6]
Institutions respond by developing data policies, governance committees, and architectural standards. Yet supervisory findings frequently arise not from the absence of policies, but from the inability to demonstrate that those policies are operating effectively at the level of specific models and decisions.
This creates a governance gap:

- Governance exists in policy documents.
- Controls exist in engineering systems.
- Assurance exists in audits and validations.

But these layers are often disconnected. The result is that data governance becomes declarative rather than demonstrable.

1.3. Reframing Data Governance as an Assurance Problem

Traditional data governance in financial institutions has evolved primarily as a managerial and organizational discipline. It focuses on defining ownership, stewardship roles, data standards, escalation processes, and policy compliance. While these elements are necessary, they are not sufficient for AI/ML systems whose behavior is highly sensitive to subtle properties of data and whose outputs can have material financial, legal, and societal consequences.
In this context, governance cannot rely solely on the existence of policies or the proper functioning of organizational processes. It must also provide independent, verifiable assurance that data-related risks are identified, controlled, and monitored in a way that is both technically sound and regulatorily defensible. [1, 2, 5]
This motivates a shift from governance as a descriptive construct ("what should happen") to governance as an operational and evidentiary construct ("what can be demonstrated to have happened").
Reframing data governance as an assurance problem implies that every material data-related claim — for example, that training data are representative, that transformations are documented and approved, that sensitive attributes are protected, or that drift is monitored — must be supported by concrete artifacts, metrics, and test procedures. These artifacts must be:

- Persistent, so that historical states can be reconstructed,
- Tamper-evident, so that integrity can be verified,
- Model-linked, so that data controls can be tied to specific model uses,
- Auditable, so that independent reviewers can assess their adequacy.

Formally, for any model output $y_t = f_{\theta_t}(x_t)$, assurance requires the ability to reconstruct and verify the full provenance of that output, including the dataset version, transformation logic, governance controls, and monitoring results in effect at time t. This can be represented by the provenance tuple:

$$\Pi_t = (D_t, T_t, G_t, M_t),$$

where D_t denotes the dataset state, T_t the transformation pipeline, G_t the governance controls applied, and M_t the monitoring and review processes in place. Governance becomes meaningful not when these elements are defined, but when Π_t can be reliably reconstructed and validated for any material decision.
To transition from a descriptive governance construct to a computable assurance mechanism, the T-ETL architecture operationalizes each element of the provenance tuple into an audit-ready state. Specifically, the dataset state D_t is verified through unique cryptographic hash commitments that provide tamper-evident integrity guarantees. The transformation pipeline T_t is represented as a versioned Directed Acyclic Graph (DAG), enabling automated lineage reconstruction and dependency analysis. Governance controls G_t are encoded as executable policy predicates—such as data quality, representativeness, and bias thresholds—that must be satisfied for model training or deployment. Finally, monitoring M_t consists of continuous telemetry over data stability and distributional metrics, supporting the timely detection of drift and other forms of input instability. Together, these operationalizations define the conditions under which provenance becomes independently verifiable rather than

merely documented.

This assurance perspective aligns closely with supervisory expectations of "effective challenge" and independent validation. It provides a bridge between technical implementation and regulatory oversight, allowing institutions to demonstrate not only that controls exist, but that they operate as intended and produce verifiable evidence.

1.4. Contribution of This Paper

The contribution of this paper is not to introduce new regulatory principles or to replace existing data governance frameworks, but to integrate and operationalize them within a unified, audit-oriented structure that is specifically designed for AI/ML systems in regulated financial environments.

More precisely, the paper makes four contributions.

First, it proposes a model-centric view of data governance, in which training and production datasets are treated as first-class model artifacts rather than as generic enterprise data assets. This shift enables governance, lineage, and quality controls to be explicitly linked to specific models, decisions, and risk uses.

Second, it introduces the Transparent ETL (T-ETL) architecture as a way to embed governance and assurance mechanisms directly into data pipelines. By integrating lineage capture, policy enforcement, and integrity verification into ingestion, transformation, and deployment workflows, T-ETL reduces the gap between governance design and operational practice.

Third, it formalizes key governance concepts — such as lineage, integrity, drift, and bias — as computable and testable properties. This allows data governance to move from qualitative assertions to quantitative evaluation, supporting repeatable audits and consistent supervisory review.

Fourth, it maps these technical and operational mechanisms explicitly to regulatory expectations under SR 11-7, BCBS 239, the EU AI Act, and DORA. This mapping translates abstract regulatory language into concrete control objectives, audit tests, and evidence artifacts, enabling institutions to demonstrate compliance in a structured and defensible manner. In doing so, the framework provides a novel assurance linkage between high-level regulatory expectations and the low-level technical evidence required for independent verification.

Taken together, these contributions position data governance not as a peripheral compliance function, but as a core component of AI/ML risk assurance. The framework is intended to be descriptive rather than prescriptive, offering a reference structure that institutions can adapt to their specific regulatory context, technological stack, and risk profile.

2. The Regulatory Imperative: Evolution of Data Mandates

The move toward verifiable data governance is driven by a convergence of global regulatory standards that, while originating from different jurisdictions and focusing on different risk domains, collectively define the requirements for AI-ready data.

2.1. The Convergence of Model, Data, and Technology Risk

Historically, regulatory oversight in financial institutions treated model risk, data risk, and technology risk as related but distinct domains. Model risk management focused on the conceptual soundness and performance of quantitative models, data governance addressed the quality and ownership of enterprise data assets, and technology risk concentrated on system availability, security, and resilience.

The increasing deployment of AI/ML systems has blurred these distinctions. Modern models do not merely process data; they internalize it through statistical learning. As a result, deficiencies in data quality, provenance, or governance propagate directly into model behavior, making it increasingly difficult to separate data risk from model risk or operational risk. [3, 4]

Regulators have responded to this convergence by expanding the scope of governance expectations. Supervisory attention has shifted from isolated control functions toward integrated, end-to-end assurance across the full data-model-decision chain. This evolution is reflected in the alignment between model risk guidance (such as SR 11-7), data aggregation principles (BCBS 239), AI-specific regulation (EU AI Act), and operational resilience requirements (DORA). While these frameworks originate from different policy concerns, they collectively articulate a consistent set of expectations: financial institutions must be able to explain, evidence, and defend the behavior of AI-driven systems under both normal and stressed conditions. [1, 2, 5, 6]

2.2. SR 11-7 and the Treatment of Data as a Model Component

Supervisory Regulation 11-7 establishes the foundational expectations for model risk management in financial institutions by defining a model as a quantitative method that transforms input data into estimates, predictions, or classifications. Within this definition, data is not treated as an external dependency, but as an intrinsic component of the model itself. This conceptualization is particularly important for AI and machine learning systems, whose behavior is largely determined by the statistical properties of the data on which they are trained. [1]

Under SR 11-7, institutions are required to demonstrate the conceptual soundness of their models, which includes the suitability, relevance, and quality of input data. For traditional parametric models, this assessment focused primarily on whether the data satisfied the assumptions of the model specification. In contrast, for AI/ML models, conceptual soundness depends on whether the training data is representative of the intended population, whether key features are stable over time, and whether data preprocessing choices introduce bias or distort the underlying signal.

SR 11-7 also requires ongoing monitoring of model performance and limitations. In an AI context, this extends beyond tracking predictive accuracy to monitoring changes in the statistical properties of input data and their relationship to model outputs. Shifts in data distributions, changes in upstream data sources, or modifications to

preprocessing pipelines can all materially alter model behavior even when the model parameters themselves remain unchanged. From a risk perspective, this means that data drift and data quality degradation become forms of model risk.

Finally, SR 11-7 emphasizes the role of effective challenge and independent review. For AI/ML systems, effective challenge cannot be limited to reviewing model code or performance metrics; it must also include scrutiny of data provenance, data preparation processes, and the assumptions embedded in the construction of training datasets. This requires institutions to maintain sufficient documentation, lineage, and evidence to allow independent parties to assess whether data-related risks are being appropriately identified, measured, and controlled

2.3. BCBS 239 and the Requirement for Traceable, Aggregatable Data

The Basel Committee's BCBS 239 framework was introduced to ensure that financial institutions have the ability to identify, aggregate, and report risk exposures in a timely, accurate, and comprehensive manner. Although BCBS 239 was not developed specifically for AI systems, its principles have direct relevance for AI/ML models that rely on complex, multi-source data pipelines to produce risk-relevant outputs. [2]

BCBS 239 requires that risk data be accurate, complete, timely, and adaptable, and that institutions maintain clear traceability from reported risk figures back to their underlying data sources. In practice, this means that institutions must be able to demonstrate how data flows from source systems through transformation processes into risk reports and decision systems. For AI-driven models, this requirement extends beyond traditional reporting into the model development and deployment lifecycle itself.

As AI models increasingly inform credit decisions, stress testing, liquidity analysis, and market risk measurement, the data feeding these models becomes part of the institution's risk data aggregation capability. If data lineage is incomplete or opaque, institutions cannot reliably explain how a given output was produced or how changes in upstream data propagate into downstream risk metrics. This undermines both internal risk management and external supervisory review.

BCBS 239 therefore implies a need for end-to-end, granular data lineage that links source data, transformation logic, model inputs, and reported outputs into a coherent and auditable chain. This chain must be sufficiently detailed to support both routine risk reporting and exceptional analysis during periods of stress or supervisory inquiry. In this sense, BCBS 239 provides the regulatory foundation for treating data traceability as a core element of AI model governance rather than as a purely technical or operational concern.

2.4. The EU AI Act and Formal Data Governance Obligations

The European Union's Artificial Intelligence Act represents a shift from principles-based guidance toward binding legal obligations for the development and use of high-risk AI systems. In the context of financial services, systems used for creditworthiness assessment, insurance risk evaluation, fraud detection, and similar decision-making functions are explicitly classified as high-risk and are therefore subject to enhanced governance and compliance requirements. [5]

A central feature of the EU AI Act is the formalization of data governance as a regulatory obligation rather than a best practice. Article 10 of the Act specifies that training, validation, and testing datasets must be relevant, representative, free of errors, and complete in relation to the intended purpose of the system. This requirement reflects the recognition that data quality and representativeness are fundamental determinants of model behavior and risk.

The Act also requires institutions to document the origin of data, the methods used to collect and preprocess it, and the assumptions made about what the data is intended to represent. This documentation must be sufficiently detailed to enable external review and supervisory assessment. As a result, data governance is no longer solely an internal control function but becomes part of the institution's regulatory compliance posture.

In addition, the Act introduces explicit expectations regarding the identification and mitigation of bias. Institutions must assess whether training data contains systematic distortions or gaps that could lead to discriminatory outcomes and must implement measures to detect and address such issues. This shifts bias from an abstract ethical concern into a concrete risk category that must be managed, monitored, and evidenced.

Taken together, these provisions establish data governance as a legally enforceable component of AI risk management. Institutions must therefore be able to demonstrate not only that appropriate data governance policies exist, but that they are implemented in practice, monitored over time, and supported by verifiable records that can be presented to regulators or other external stakeholders.

2.5. DORA and the Operational Resilience of Data

The Digital Operational Resilience Act (DORA) extends regulatory oversight beyond models and data content to the operational infrastructure that supports digital financial services. While DORA is often discussed in the context of cybersecurity and ICT risk, its provisions have direct implications for data governance in AI/ML systems, particularly with respect to data integrity, availability, and recoverability. [6]

DORA requires financial institutions to identify and manage their critical ICT dependencies, maintain detailed records of digital operations, and implement controls to ensure the continuity and reliability of information systems. For AI-driven processes, this includes the systems that ingest, transform, store, and serve training and production data. As such, data pipelines become part of the institution's operational resilience framework.

A key implication of DORA for data governance is the requirement that institutions be able to recover and verify the integrity of critical data assets following operational disruptions. This includes cyber incidents, system outages, or failures at third-party service providers. In the context of AI/ML, institutions must therefore ensure that training datasets, model inputs, and associated metadata can be restored to a known and trusted state after such events.

DORA also emphasizes the need for audit-ready logging and traceability of ICT activities. For data governance, this

implies that data processing events, changes to data pipelines, and access to sensitive datasets must be recorded in a manner that supports forensic analysis and supervisory review. These requirements reinforce the need for persistent, tamper-evident records that link operational events to data states and model behavior.

In this way, DORA integrates operational resilience into the broader governance of AI/ML systems. Data governance is no longer only about quality and appropriateness, but also about ensuring that data remains trustworthy and available under adverse conditions. This resilience perspective complements the governance and risk management requirements of SR 11-7, BCBS 239, and the EU AI Act, completing the regulatory foundation for auditable and reliable AI-driven decision systems in financial institutions.

3. Reframing Data Governance as an Assurance Problem

The regulatory developments discussed in the previous section establish that data governance for AI/ML systems in financial institutions is no longer a matter of internal policy or technical preference. It is a regulated, auditable, and enforceable requirement. However, while regulations specify what must be achieved — transparency, traceability, integrity, representativeness, and resilience — they provide limited guidance on how these objectives should be operationalized within complex, dynamic AI data pipelines.

This gap motivates a conceptual shift: data governance must be reframed from a management discipline into an assurance discipline. That is, governance should not be defined solely by organizational structures, policies, or standards, but by the institution's ability to provide independent, verifiable evidence that data-related risks are being appropriately identified, controlled, and monitored over time.

Under this framing, the central question of data governance becomes not "Are appropriate policies in place?" but "Can the institution demonstrate, at any point in time, how a specific model decision was produced, what data it relied on, what transformations were applied, what controls were enforced, and how risks were monitored?"

This reframing shifts the focus from static governance artifacts to dynamic, verifiable processes.

3.1. Datasets as Controlled Model Artifacts

A foundational principle of the proposed framework is that training and production datasets should be treated as controlled model artifacts rather than as generic enterprise data assets.

In traditional data management, data is often governed at the level of domains, systems, or business functions. In AI/ML systems, however, data acquires model-specific meaning: the same dataset can have different risk implications depending on how it is used, which model consumes it, and what decisions it informs.

Treating data as a controlled model artifact implies that datasets are:

- Explicitly versioned,
- Persistently stored in immutable form once used for training or validation,
- Linked to specific model versions and uses,
- Subject to approval, review, and change control processes.

This ensures that models are reproducible not only in terms of code and parameters, but also in terms of the data on which they were trained and evaluated. It also enables institutions to demonstrate that changes in model behavior are attributable to known changes in data or model logic, rather than to undocumented or uncontrolled factors.

From a risk perspective, this approach acknowledges that in data-driven systems, data functions as a form of executable logic. Just as changes to model code can alter outcomes, so can changes to the data on which models are trained or operate. Governing data as a controlled artifact therefore extends core principles of software and model governance into the data domain.

3.2. The Assurance Perspective

From an assurance perspective, governance claims must be testable. Assertions such as "training data is representative," "transformations are documented," or "drift is monitored" only become meaningful when they can be independently verified.

The framework therefore emphasizes the creation and preservation of evidence artifacts that support such verification. These artifacts include:

- Versioned datasets and associated metadata,
- Documented transformation and feature engineering logic,
- Lineage representations linking source data to model inputs and outputs,
- Logs of access, modification, and processing events,
- Records of quality checks, bias assessments, and drift monitoring.

These artifacts serve two purposes. Internally, they support effective challenge by enabling model validators, risk managers, and auditors to understand and test how models and data behave. Externally, they support regulatory and supervisory review by providing concrete evidence of compliance with governance expectations.

The assurance perspective also changes the role of governance functions. Rather than primarily setting policies and standards, governance becomes responsible for defining what constitutes acceptable evidence, how that evidence is produced, and how it is reviewed and maintained over time.

An assurance-oriented approach also implies repeatability and independence. Governance evidence must be generated in a consistent manner across models and over time, such that different reviewers, operating independently, would reach materially similar conclusions when examining the same artifacts. This distinguishes assurance from compliance reporting, which often relies on narrative descriptions or one-time attestations. In the proposed framework, assurance is achieved when data governance claims can be validated without reliance on discretionary explanations by model developers or data engineers, but instead through standardized, preserved, and independently reviewable evidence.

3.3. The Provenance Chain

At the core of the framework is the concept of a provenance chain that links model decisions back to their underlying data and governance context.

For any model decision made at time ttt, the institution should be able to reconstruct:

- The dataset state in effect at that time,
- The transformation and feature engineering logic applied,
- The governance controls enforced,
- The monitoring and review processes active.

Together, these elements form a provenance chain that supports both explainability and accountability.

This chain enables several critical capabilities:

- Reproducibility — the ability to recreate model behavior under historical conditions,
- Explainability — the ability to articulate how specific data contributed to a given outcome,
- Forensic analysis — the ability to investigate incidents, errors, or complaints,
- Regulatory defense — the ability to provide coherent evidence during supervisory review.

Without such a chain, institutions may be able to describe their governance structures, but not to demonstrate their operation in specific cases.

3.4. Overview of the Framework

The proposed framework operationalizes this assurance perspective through a set of integrated components that span the data lifecycle:

1. A data ingestion layer that captures provenance at the point of entry,
2. A metadata and lineage layer that records transformations and dependencies,
3. A governance and control layer that enforces policies and quality thresholds,
4. A monitoring layer that detects drift, bias, and anomalies,
5. An evidence and audit layer that preserves artifacts for review and reconstruction.

These components are not independent. They are designed to function as a coherent system in which governance requirements are embedded into technical processes and technical artifacts are structured to support governance and assurance.

This integration distinguishes the framework from approaches that treat governance as an overlay on existing systems. Instead, governance becomes a property of the system's design.

3.5. Positioning Relative to Existing Work

Existing literature addresses data governance, lineage, model risk, and AI ethics as related but largely separate domains. Governance frameworks often emphasize organizational roles and policies, while technical work focuses on specific mechanisms such as lineage capture, drift detection, or bias measurement.

This paper does not attempt to replace or subsume these strands of work. Rather, it integrates them within a model-centric, assurance-oriented structure that is explicitly aligned with regulatory expectations and audit practice.

In this sense, the contribution of the framework is not the invention of new governance principles, but the translation of existing principles into an operational and evidentiary form suitable for regulated AI systems.

Operational lineage tools such as OpenLineage and Marquez provide metadata capture and visualization of pipeline dependencies; however, they do not establish a cryptographic root-of-trust at ingestion, encode governance controls as executable policy predicates, or produce deterministic auditor test procedures mapped to supervisory mandates. The T-ETL architecture extends beyond passive lineage capture by functioning as a verification engine whose primary output is audit-grade provenance evidence.

4. The Transparent ETL (T-ETL) Architecture

The assurance-oriented framework introduced in the previous section requires a technical substrate capable of capturing, preserving, and exposing governance-relevant information about data as it moves through its lifecycle. Traditional extract-transform-load (ETL) pipelines are optimized for efficiency, scalability, and reliability, but they are not designed to support auditability, reproducibility, or regulatory review.

In conventional pipelines, data transformations are often implicit in code, metadata is incomplete or transient, and historical states are overwritten. This makes it difficult, and often impossible, to reconstruct how a particular model input was produced, what assumptions were embedded in preprocessing, or how changes in upstream data influenced downstream decisions.

To address this structural limitation, this paper proposes the Transparent ETL (T-ETL) architecture — an extension of conventional data pipelines that embeds provenance capture, governance enforcement, and evidence preservation directly into data ingestion, transformation, and deployment processes.

The objective of T-ETL is not to replace existing data infrastructure, but to introduce a parallel, governance-aware layer that makes data behavior observable, traceable, and defensible.

Unlike MLOps (machine learning operations frameworks focused on deployment and performance monitoring) and data observability frameworks, which primarily optimize operational reliability, deployment efficiency, and performance monitoring, T-ETL is explicitly designed to generate audit-grade evidence, preserve historical data states, and enable independent reconstruction of model decisions under regulatory review.

4.1. Design Principles

The design of the Transparent ETL (T-ETL) architecture is guided by five core principles: transparency, persistence, immutability, linkage, and verifiability. Transparency ensures that data flows and transformations are observable and inspectable rather than implicit or opaque. Persistence and immutability ensure that governance-relevant artifacts are retained over time and protected from retroactive alteration, enabling historical reconstruction and auditability. Linkage connects datasets explicitly to model versions, decisions, and business uses so that governance follows impact rather than organizational boundaries. Verifiability requires that all

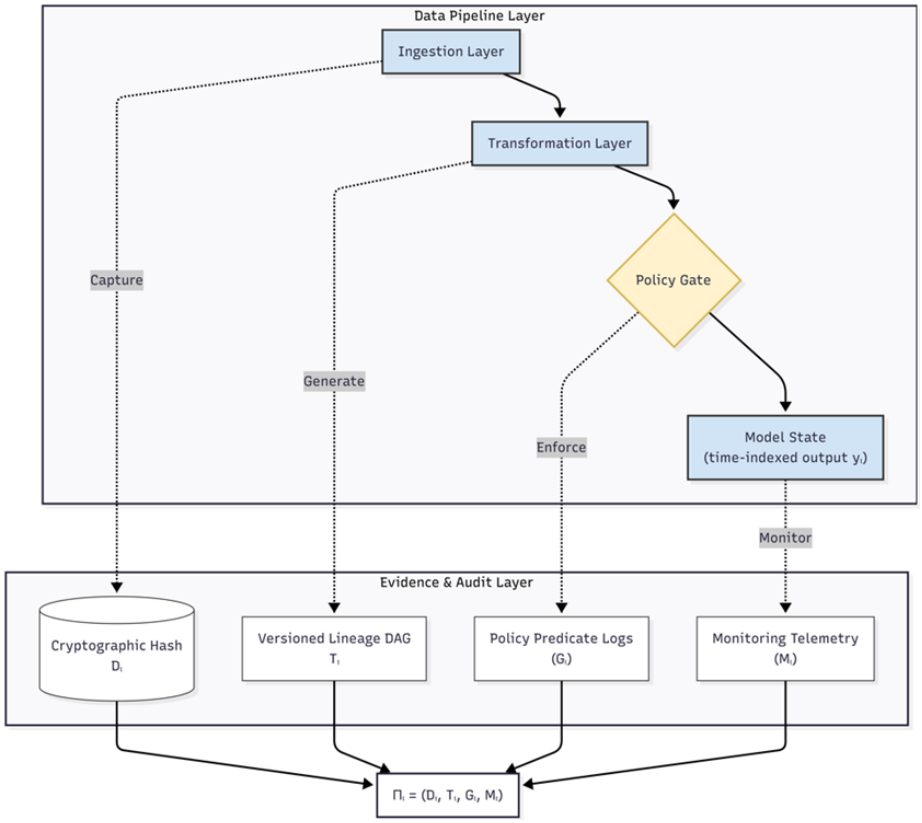
governance claims be supported by testable evidence, specifically, the architecture must automatically generate, persist and version the components of the provenance tuple ΠIt at each stage of the data and model lifecycle.
Under this design, T-ETL functions not merely as a data processing pipeline, but as a verification engine whose primary output is audit-grade evidence. Each ingestion, transformation, governance check, and monitoring event deterministically produces verifiable artifacts corresponding to the state of Dt, Tt, Gt, and Mt at that point in time. As a result, provenance is not reconstructed retrospectively for audit purposes, but is continuously and automatically asserted as part of normal system operation. Together, these principles align technical system design with regulatory and assurance objectives. They ensure that governance is not an overlay applied after the fact, but a property embedded into how data is ingested, transformed, monitored, and retained. As a result, the data pipeline itself becomes a source of assurance, continuously generating the artifacts required for effective challenge, audit, and regulatory review.

4.2. Architecture Overview

T-ETL introduces five tightly integrated layers that extend traditional data pipelines:

1. **Ingestion Layer:** Captures raw data and contextual metadata.
2. **Transformation Layer:** Applies and documents preprocessing and feature engineering.
3. **Lineage and Metadata Layer:** Records dependencies between all artifacts.
4. **Governance and Control Layer:** Enforces policy and risk constraints.
5. **Evidence and Audit Layer:** Preserves artifacts for reconstruction and review.

Each layer both produces and consumes governance artifacts, forming a closed assurance loop. Figure 1 provides a system-level view of the Transparent ETL (T-ETL) architecture and its role as a provenance verification engine. The upper layer represents the operational data pipeline, while the lower layer captures the audit and evidence substrate generated at each stage. Each pipeline component deterministically produces a corresponding element of the provenance tuple—dataset state (D_t), transformation lineage (T_t), governance enforcement (G_t), and monitoring telemetry (M_t). By explicitly separating operational processing from evidence persistence, the architecture ensures that provenance is continuously asserted during normal operation rather than reconstructed retrospectively for audit purposes.



System-level architecture of the Transparent ETL (T-ETL) framework. The upper layer represents the operational data pipeline, including ingestion, transformation, and policy enforcement. The lower Evidence and Audit Layer persistently captures cryptographic dataset hashes (D_t), versioned lineage graphs (T_t), executable governance predicates (G_t), and monitoring telemetry (M_t). Together, these artifacts compose the provenance tuple Π_t, enabling audit-ready reconstruction, integrity verification, and regulatory defensibility of AI/ML model decisions.

Fig 1: Transparent ETL (T-ETL) Architecture as a Provenance Verification Engine

4.3. Ingestion Layer

The Ingestion Layer acts as the root of the chain of trust. Upon data capture, it generates a cryptographic hash commitment for each raw dataset version, thereby establishing the immutable Dt component of the provenance tuple. This commitment ensures that any subsequent unauthorized alteration to source data is detectable during audit or forensic review.

The ingestion layer establishes the authoritative entry point for all data used in modeling.

Beyond simple extraction, it performs:

- Source identification and authentication,
- Schema validation and consistency checks,
- Capture of extraction time, system state, and responsible owner,
- Classification of data sensitivity and regulatory relevance.

It produces:

- Immutable raw data snapshots,
- Source provenance metadata,
- Initial data quality indicators.

These outputs enable auditors to verify where data originated, under what conditions it was collected, and whether it was appropriate for its intended use.

4.4. Transformation Layer

Each transformation step is treated as a node in a versioned Directed Acyclic Graph (DAG). By capturing the exact transformation logic Δ applied between nodes, this layer constructs the Tt, component of the provenance tuple, allowing auditors to traverse the graph and verify how raw inputs were mathematically mapped into model-ready features.

The transformation layer is responsible for converting raw data into model-ready inputs.

Each transformation step is:

- Explicitly declared,
- Parameterized,
- Version-controlled,
- Linked to justification and approval records.

The layer records:

- Input and output dataset identifiers,
- Transformation logic versions,
- Feature definitions and derivations,
- Known limitations and assumptions.

This enables independent reviewers to assess whether transformations introduce bias, leakage, or instability.

4.5. Lineage and Metadata Layer

The lineage layer represents the entire pipeline as a directed acyclic graph, with nodes representing datasets, transformations, and models, and edges representing dependencies.

This supports:

- Forward impact analysis,
- Backward root-cause analysis,
- Regulatory reconstruction of decisions,
- Scenario analysis for change management.

The graph structure enables both technical diagnostics and governance oversight.

4.6. Governance and Control Layer

This layer encodes governance policies as executable controls. This layer evaluates executable policy predicates—computable gates for data quality, bias, and representativeness. The outcomes of these evaluations are persisted as the Gt component of the provenance tuple, providing an explicit approval trail that demonstrates governance enforcement rather than mere policy documentation.

Controls may include:

- Minimum data quality thresholds,
- Bias and representativeness checks,
- Access restrictions,
- Approval workflows for changes.

Controls can be enforced automatically or routed for human review.

The results of control evaluations are stored as evidence.

4.7. Evidence and Audit Layer

The Evidence and Audit Layer serves as the system of record for the provenance tuple Πt. It indexes dataset hash commitments, lineage DAGs, and policy evaluation logs into an immutable, append-only store, providing a unified reference point for forensic reconstruction, independent audit, and regulatory review:

- Dataset versions,
- Lineage graphs,
- Control outcomes,
- Access and change logs,
- Monitoring results.

Artifacts are stored immutably and indexed for retrieval. This enables reconstruction of historical states and supports audits, regulatory reviews, and internal investigations

4.8. Summary

T-ETL transforms data pipelines into auditable systems. It provides the technical infrastructure necessary to support assurance-oriented data governance and allows institutions to demonstrate compliance, accountability, and control.

5. Audit Tests, Metrics, and Evidence for Data Governance

The assurance-oriented framework and T-ETL architecture described in the previous sections are only meaningful if they can be independently tested and verified. Governance, in a regulatory context, is not assessed by design intent but by demonstrable effectiveness. This section translates the proposed framework into concrete audit tests, metrics, and evidence artifacts that enable institutions to assess and defend the integrity, traceability, and reliability of data used in AI/ML systems.

Rather than prescribing specific technologies or tools, this section focuses on control objectives and testable outcomes. These tests are designed to support internal

audit, independent model validation, and supervisory review.

A worked, audit-oriented illustration of these verification procedures is provided in Appendix B, demonstrating how the provenance tuple Π_t can be independently reconstructed and validated using preserved evidence artifacts.

5.1. Data Integrity and Quality Verification

Data integrity and quality form the foundation of AI/ML model reliability. Audit tests in this area assess whether data used for training and production is complete, accurate, consistent, and protected from unauthorized alteration.

In formal terms, these controls are designed to ensure that the empirical input distribution P(X) used by the model fθ remains consistent with the governed dataset state Dt captured in the provenance tuple Πt, thereby preventing unmonitored changes in data from propagating into model behavior.

Audit Test: Cryptographic Integrity Verification (Dt) To verify data integrity, an auditor may retrieve the cryptographic hash commitment associated with the dataset version Dt from the Evidence and Audit Layer and independently re-calculate the hash over the archived dataset. A successful match provides mathematical assurance that the training or production data has not been subject to unauthorized alteration since the point of ingestion. This test transforms data integrity from a qualitative assertion into a computable, bit-level verification.

Key audit objectives include:

- Verifying that datasets used for model development and deployment are versioned and immutable once approved for use.
- Confirming that data completeness thresholds are defined and monitored.
- Assessing whether data validation checks are applied consistently at ingestion and transformation stages.

Evidence artifacts may include dataset snapshots, validation reports, schema checks, and integrity logs. Auditors should be able to trace quality issues to specific data sources or pipeline stages and assess whether remediation actions were appropriately taken.

5.2. Lineage and Traceability Verification

Lineage verification assesses whether institutions can demonstrate end-to-end traceability from model outputs back to underlying data sources and transformations.

Audit tests in this area evaluate:

- Whether lineage information is captured automatically rather than reconstructed manually.
- Whether lineage is sufficiently granular to support attribute-level tracing.
- Whether lineage links datasets, transformations, and model versions in a consistent and reproducible manner.

Audit Test: Algorithmic Traceability (Tt) Leveraging the versioned Directed Acyclic Graph (DAG), the auditor may perform a retrospective reconstruction on a selected model output. By traversing the edges of the DAG, the auditor verifies that the transformation logic Δ applied to the raw inputs corresponds to the approved transformation definitions and documented model design. This test demonstrates that feature engineering remains transparent, reproducible, and traceable throughout the model lifecycle.

Auditors should be able to select a model output or risk metric and reconstruct the full data path that produced it, including intermediate transformations and feature engineering steps. The absence of complete lineage represents a material governance weakness, particularly for high-impact models.

5.3. Drift, Bias, and Stability Monitoring

AI/ML models are sensitive to changes in data distributions and relationships over time. Audit tests therefore assess whether institutions monitor and respond to data drift, bias, and instability in a systematic manner.

Key questions include:

- Are baseline data distributions documented and preserved?
- Are changes in input data characteristics monitored over time?
- Are thresholds defined for material deviations, and are breaches escalated appropriately?

Evidence may include drift reports, bias assessments, monitoring dashboards, and escalation records. Auditors should assess not only whether monitoring exists, but whether it is linked to decision-making and governance actions, such as model review or retraining.

Audit Test: Policy Predicate Enforcement (Gt) The auditor may inspect policy evaluation logs to confirm that each dataset version used in training or production was evaluated against—and satisfied—the defined bias, representativeness, and quality predicates (Γ) prior to model execution. This verification provides evidence that governance controls functioned as enforceable gates rather than post-hoc documentation, supporting assurance expectations for high-risk AI systems.

Audit Test: Monitoring Telemetry Verification (Mt) To verify the monitoring component of the provenance tuple, the auditor may inspect preserved telemetry records to confirm that input distribution statistics, stability metrics, and drift indicators were continuously captured for the governed dataset state Dt. This test ensures that material deviations in the empirical input distribution P(X),) would have been observable and subject to escalation, completing the verification of the provenance Πt = (Dt,Tt,Gt,Mt).

5.4. Control Effectiveness and Exception Management

This subsection evaluates whether governance controls embedded in the data pipeline operate effectively in practice.

Audit tests focus on:

- Whether data governance controls are enforced automatically or rely on manual intervention.
- How exceptions are identified, documented, and approved.
- Whether control breaches are tracked and resolved within defined timelines.

Effective governance requires not only strong controls but also transparent handling of exceptions. Evidence should demonstrate that deviations from policy are intentional, justified, and subject to appropriate oversight rather than ad hoc or undocumented.

5.5. Evidence Sufficiency and Audit Readiness

Finally, auditors assess whether the institution's data governance framework produces sufficient and reliable evidence to support regulatory review.

This includes evaluating:

- The completeness and retention of governance artifacts.
- The accessibility of evidence for independent review.
- The consistency of evidence across models and data pipelines.

Audit readiness is achieved when an institution can respond to inquiries with concrete artifacts rather than narrative explanations. This capability is central to demonstrating effective data governance under modern regulatory expectations.

Appendix B demonstrates an end-to-end execution of the cryptographic integrity, lineage reconstruction, and policy predicate enforcement tests on a synthetic dataset, enabling independent reproduction of the verification steps described in this section.

6. Governance, Roles, and Operating Model

The assurance-oriented data governance framework described in the previous sections requires a governance and operating model that clearly defines accountability, independence, and escalation pathways. In regulated financial institutions, governance effectiveness depends not only on technical controls, but also on how responsibilities are allocated, how decisions are challenged, and how issues are resolved across organizational boundaries.

This section describes a governance and operating model that aligns data governance for AI/ML systems with existing risk management, model governance, and audit structures, while addressing the unique challenges posed by data-driven models.

6.1. Roles and Responsibilities

Effective data governance for AI/ML systems requires clearly defined roles across the data, model, risk, and assurance functions. Ambiguity in ownership or accountability is a common root cause of governance failures.

Key roles typically include:

- Data Owners, who are accountable for the quality, appropriateness, and permitted use of data sources.
- Data Stewards, who oversee day-to-day data management, metadata maintenance, and issue remediation.
- Model Owners, who are accountable for model performance, limitations, and compliance with governance requirements.
- Model Validators, who provide independent assessment of model design, data dependencies, and monitoring practices.
- Risk Management Functions, which set risk appetite, define governance standards, and oversee adherence.
- Internal Audit, which provides independent assurance over the effectiveness of governance controls.

In the proposed framework, these roles are connected through shared access to governance artifacts generated by the T-ETL architecture, reducing reliance on informal communication and manual documentation.

6.2. Integration with Model Risk Management

Data governance for AI/ML systems should not operate as a standalone function. It must be integrated into the institution's broader model risk management (MRM) framework.

This integration includes:

- Treating data-related risks as explicit components of model risk assessments.
- Incorporating data lineage, quality, and drift evidence into model validation reviews.
- Aligning data governance approval and change management processes with model lifecycle stages.

By embedding data governance into MRM, institutions ensure that data risks are evaluated with the same rigor as model assumptions and performance metrics.

A critical element of this integration is the preservation of independence and separation of duties. While data engineers and model developers are responsible for implementing data pipelines and controls, assurance requires that the adequacy of these controls be assessed by functions that are not directly involved in their design or operation. By generating standardized governance artifacts through the T-ETL architecture, the framework enables independent model validation, risk management, and internal audit functions to evaluate data-related risks without reliance on self-attestation by model owners. This supports effective challenge and reinforces accountability across the model lifecycle

6.3. Alignment with Operational and Technology Risk

AI/ML data pipelines depend heavily on technology infrastructure and third-party services. As a result, data governance must also align with operational risk and technology risk management.

This includes:

- Mapping data pipelines to critical ICT assets and dependencies.
- Ensuring that data governance controls are considered in operational resilience planning.
- Coordinating incident response procedures for data-related events.

This alignment helps prevent gaps where data risks fall between organizational silos

6.4. Decision-Making and Escalation Processes

An effective governance model requires clear decision-making and escalation mechanisms.

Institutions should define:

- Criteria for identifying material data governance issues.
- Thresholds for escalation to senior management or governance committees.

- Procedures for approving exceptions or deviations from standards.

Escalation processes should be evidence-based, relying on artifacts produced by the governance framework rather than subjective judgment.

6.5. Governance Committees and Oversight

Oversight bodies such as model risk committees, data governance councils, or risk committees play a central role in setting expectations and resolving cross-functional issues.

In the context of AI/ML data governance, these bodies should:

- Review material data risks and trends.
- Approve significant changes to data pipelines or governance controls.
- Ensure alignment between business objectives, risk appetite, and regulatory requirements.

Committee oversight ensures that governance decisions reflect institutional priorities and risk tolerance

6.6. Continuous Improvement and Accountability

Finally, the governance model must support continuous improvement. As AI/ML systems evolve, so do data sources, risks, and regulatory expectations.

Institutions should periodically review:

- The effectiveness of data governance controls.
- The adequacy of evidence and audit artifacts.
- Lessons learned from incidents, audits, or regulatory feedback.

Continuous improvement ensures that the governance framework remains relevant and effective over time.

7. Comparative Analysis and Illustrative Scenarios

The assurance-oriented data governance framework proposed in this paper differs from both traditional data governance approaches and many contemporary AI governance initiatives in its scope, focus, and operational emphasis. This section compares the proposed framework with existing practices and illustrates its application through representative scenarios that highlight common governance challenges in AI/ML systems.

7.1. Comparison with Traditional Data Governance Approaches

Traditional data governance frameworks in financial institutions are primarily designed to manage enterprise data assets across business domains. They emphasize ownership, stewardship, data definitions, and policy compliance, often focusing on static datasets used for reporting and analytics.

In contrast, AI/ML systems introduce dynamic data dependencies, frequent retraining, complex feature engineering, and continuous deployment. Traditional governance approaches struggle to address these characteristics because they are not designed to capture fine-grained lineage, transformation logic, or time-specific data states.

The proposed framework differs in three key respects:

- **Model-centric focus:** Data governance is explicitly linked to model usage and decision impact rather than organized solely around data domains.
- **Assurance orientation:** Governance effectiveness is evaluated based on verifiable evidence and auditability rather than policy existence.
- **Lifecycle integration:** Governance controls are embedded into data pipelines across ingestion, transformation, monitoring, and retention.

These differences enable the framework to address governance risks that are otherwise difficult to manage in AI-driven environment

7.2. Comparison with Existing AI Governance Frameworks

Many AI governance frameworks emphasize ethical principles, transparency objectives, or high-level risk management guidelines. While these approaches provide valuable direction, they often lack concrete mechanisms for operational enforcement and independent verification.

The framework proposed in this paper complements such initiatives by focusing on the data layer as a primary source of AI risk. Rather than defining new ethical principles, it operationalizes existing regulatory expectations by translating them into auditable technical and organizational controls.

This approach shifts governance discussions from abstract concepts to concrete questions such as:

- Can the institution reconstruct the exact data used to train a model?
- Can it demonstrate how data transformations affected model inputs?
- Can it evidence how data risks were monitored and addressed over time?

By answering these questions, institutions can bridge the gap between governance intent and demonstrable compliance.

7.3. Illustrative Scenario: Data Drift in a Credit Risk Model

Consider a credit risk model deployed to assess borrower default probability. Over time, changes in customer behavior, economic conditions, or data collection processes alter the statistical properties of input data.

In a traditional governance setup, such drift may go undetected until model performance degrades materially. Investigating the root cause can be time-consuming, requiring manual reconstruction of data pipelines and assumptions.

Under the proposed framework, drift detection metrics and lineage artifacts generated by the T-ETL architecture allow institutions to identify when data distributions diverge from historical baselines. Governance controls trigger escalation and review, enabling timely intervention such as model recalibration or retraining. Auditors can subsequently verify that the issue was detected, assessed, and addressed in accordance with governance standards.

Within the formal framework introduced earlier, such drift corresponds to a divergence between the observed input distribution $P(X)$ and the baseline distribution associated with the original dataset state D_t, allowing the provenance

tuple ΠIt to be used to localize and evidence the source of the deviation.

In a supervisory or crisis context, this capability allows the institution to reconstruct the full provenance tuple ΠIt = (Dt,Tt,Gt,Mt), corresponding to the period immediately preceding and following the detected drift event. By comparing successive dataset states Dt and monitored telemetry Mt, auditors and regulators can verify not only that drift occurred, but precisely when it emerged, which transformations Tt were implicated, and whether governance controls Gt were appropriately enforced. This transforms drift from a retrospective performance issue into a time-indexed, auditable risk event.

7.4. Illustrative Scenario: Bias Identification and Regulatory Review

In another scenario, a financial institution faces regulatory scrutiny regarding potential bias in an AI-driven lending decision system. Regulators request evidence demonstrating that training data was representative and that bias risks were identified and mitigated.

Without structured governance artifacts, the institution may struggle to provide clear and consistent evidence. Under the proposed framework, the institution can reconstruct the specific provenance tuple ΠIt associated with the challenged decision period, including the governed dataset state Dt, the approved transformation pipeline Tt, the executed policy predicates Gt, and the contemporaneous monitoring outputs Mt.

This capability reduces reliance on retrospective explanations and strengthens the institution's regulatory defensibility.

7.5. Illustrative Scenario: Incident Investigation and Forensic Analysis

A data integrity incident occurs due to an upstream system error that corrupts a subset of production data. The incident affects model outputs used in downstream decision-making.

Using the proposed framework, the institution can trace the impact of the corrupted data through the lineage graph, identify affected models and decisions, and assess the scope of remediation required. The evidence and audit layer supports forensic analysis by preserving historical data states and processing logs.

This structured response enables faster resolution, clearer accountability, and more effective communication with regulators and stakeholders.

From an assurance perspective, such incidents represent a stress test of the institution's ability to demonstrate control under adverse conditions. By preserving the provenance tuple ΠIt for all materially affected decisions, the framework enables regulators and internal investigators to independently reconstruct the precise data state, transformations, governance controls, and monitoring context in effect at the time of failure. This capability is critical not only for remediation, but for demonstrating systemic resilience and accountability in high-impact, nationally significant financial systems.

7.6. Limitations and Trade-Offs

While the proposed framework offers significant benefits, it also introduces trade-offs. Implementing T-ETL and assurance-oriented governance requires investment in infrastructure, process redesign, and organizational change. Increased data retention and metadata capture may raise storage and operational costs.

Institutions must balance these costs against the benefits of improved auditability, reduced regulatory risk, and enhanced trust in AI-driven systems. The framework is therefore intended to be scalable and risk-based, with governance intensity aligned to the materiality of model use.

Certain deployment paradigms, such as highly decentralized or privacy-preserving training architectures, may require adaptations of the proposed framework to maintain equivalent levels of auditability and evidentiary sufficiency

8. Trends, Gaps, and Future Directions

The governance of data for AI/ML systems in financial institutions is evolving rapidly in response to technological innovation, regulatory development, and changing risk profiles. While the framework proposed in this paper addresses current governance and assurance challenges, several trends and gaps will shape how data governance for AI systems develops over the coming years.

8.1. Emerging Technologies and Data Paradigms

Several emerging technologies are likely to further complicate data governance for AI/ML systems. These include the increased use of alternative data sources, synthetic data generation, federated learning, and privacy-preserving machine learning techniques.

Alternative and non-traditional data sources introduce new questions around provenance, representativeness, and legality of use. Synthetic data, while offering potential benefits for privacy and data scarcity, raises governance challenges related to traceability, validation, and the risk of amplifying embedded biases. Federated learning and distributed training architectures reduce centralized data aggregation but complicate lineage, auditability, and assurance, as training data may never be directly observable by the model owner.

These developments reinforce the need for governance frameworks that are adaptable and capable of capturing assurance-relevant information even when data does not follow traditional centralized pipeline

8.2. Regulatory Evolution and Supervisory Expectations

Regulatory expectations for AI and data governance are likely to continue converging across jurisdictions. While current frameworks such as SR 11-7, BCBS 239, the EU AI Act, and DORA originate from different regulatory objectives, they increasingly emphasize common themes: accountability, transparency, traceability, and resilience.

Future supervisory guidance is likely to place greater emphasis on demonstrable outcomes rather than policy intent. Institutions may be expected not only to describe their governance frameworks, but to produce concrete evidence showing how data risks are identified, monitored, and controlled in practice. This trend will further elevate the importance of assurance-oriented governance and auditable technical architectures.

As supervisory scrutiny intensifies, institutions that rely on manual documentation or ad hoc lineage reconstruction

may face increasing challenges in demonstrating compliance.

8.3. Gaps in Research and Practice

Despite growing attention to AI governance, significant gaps remain in both academic research and industry practice. Much of the existing literature focuses on ethical principles, algorithmic transparency, or model explainability, with comparatively less attention to the operational governance of data across the full model lifecycle.

In practice, many institutions continue to treat data governance, model governance, and technology risk as separate domains, leading to fragmented controls and inconsistent assurance. There is limited consensus on how to operationalize data governance in environments characterized by continuous deployment, frequent retraining, and third-party data dependencies.

These gaps suggest a need for further research into scalable, audit-ready data governance architectures and for practical frameworks that bridge technical implementation and regulatory oversight.

8.4. Future Directions for Data Governance in AI

Over the next five years, data governance for AI/ML systems is likely to evolve toward more automated, embedded, and risk-sensitive approaches. Governance controls may increasingly be encoded directly into data and model pipelines, reducing reliance on manual review processes.

Institutions are also likely to adopt more granular, model-specific governance structures, reflecting the varying risk profiles of different AI applications. High-impact models may be subject to enhanced governance, monitoring, and evidence retention, while lower-risk use cases may be governed more proportionately.

Finally, as AI systems become more integrated into core financial decision-making, data governance is likely to become a central element of institutional trust and accountability. Frameworks that enable institutions to demonstrate not only compliance, but also disciplined and responsible use of data, will play an increasingly important role in sustaining confidence among regulators, customers, and society at large

9. Conclusion

The increasing reliance on AI and machine learning in financial institutions has fundamentally altered the nature of data governance. Data is no longer a passive input to analytical processes, but an active determinant of model behavior, risk outcomes, and regulatory exposure. As a result, traditional approaches to data governance—focused on policy definition, ownership, and high-level controls—are insufficient to meet the demands of modern, data-driven decision systems.

This paper has argued that data governance for AI/ML systems must be reframed as an assurance problem. By shifting the focus from managerial structures to verifiable evidence, institutions can better align technical implementation with regulatory expectations for transparency, traceability, integrity, and resilience. The proposed framework treats data as a controlled model artifact and embeds governance directly into data pipelines through the Transparent ETL (T-ETL) architecture, enabling reproducibility, effective challenge, and auditability across the model lifecycle.

A central contribution of this paper is the integration of data governance, model risk management, and audit assurance into a single, coherent framework designed specifically for regulated AI systems. While prior work has addressed data quality, lineage, or AI governance in isolation, this paper operationalizes these concepts by translating regulatory principles into concrete architectural components, control objectives, and audit-grade evidence artifacts. The T-ETL architecture and the associated assurance model provide a practical mechanism for linking regulatory expectations to technical implementation and independent verification.

By mapping regulatory requirements from SR 11-7, BCBS 239, the EU AI Act, and DORA to specific governance mechanisms and audit tests, the framework enables institutions to move beyond descriptive compliance toward demonstrable accountability. Ultimately, effective data governance for AI is not established through the existence of policies or committees, but through the ability to reconstruct, explain, and defend individual model decisions using reliable and verifiable evidence. Institutions that adopt assurance-oriented data governance architectures will be better positioned to manage emerging risks, withstand regulatory scrutiny, and sustain trust in AI-driven financial decision-making as regulatory and technological landscapes continue to evolve.

Ultimately, the contribution of this framework lies in the transition from declarative governance to demonstrable assurance. By operationalizing the provenance tuple through the T-ETL architecture, financial institutions can move beyond policy assertions to provide mathematical and forensic proof of data integrity. This shift not only supports regulatory defensibility under mandates such as SR 11-7 and the EU AI Act, but advances a technically rigorous reference model for algorithmic accountability in the global financial ecosystem.

Appendix B: Worked Example of Audit-Grade Data Verification Using T-ETL

This appendix provides a concrete, audit-oriented illustration of how the Transparent ETL (T-ETL) framework operationalizes the provenance tuple

$\Pi t = (Dt, Tt, Gt, Mt)$

to support independent verification of data integrity, traceability, and governance enforcement. The example is illustrative rather than system-specific and is intended to demonstrate verifiability, not implementation.

Appendix B.1 Dataset State Verification (Dt)

Assume a supervised learning model trained on a tabular dataset derived from multiple internal banking systems.

Each dataset version approved for model use is treated as immutable and assigned a cryptographic hash commitment at ingestion.

Table 1: Dataset Version Register

Dataset ID	Source Systems	Ingestion Timestamp	Record Count	Hash Commitment (SHA-256)
D_1	Loan Core, KYC DB	2024-03-01 10:12 UTC	1,245,311	9f3c…e12a
D_2	Loan Core, KYC DB	2024-06-01 09:55 UTC	1,308,402	c7a1…9bd4

Audit Verification Test

An auditor retrieves dataset D_1 from the Evidence and Audit Layer and independently recomputes its hash. A match against the registered commitment provides mathematical proof that the dataset has not been altered since ingestion.

This transforms data integrity from a policy assertion into a bit-level verification.

Appendix B.2 Transformation and Lineage Verification (Tt)

Data transformations are represented as a versioned Directed Acyclic Graph (DAG), with each node corresponding to a deterministic transformation step.

Table 2: Transformation DAG Nodes

Node ID	Input Dataset	Transformation Logic	Output Dataset
N_1	D_1	Missing value imputation (median)	D_1'
N_2	D_1'	Feature normalization (z-score)	D_1''
N_3	D_1''	Feature encoding (one-hot)	D_1'''

Audit Verification Test

Given a selected model output, the auditor traverses the DAG backward from D_1''' to D_1, verifying that:

- All transformations were approved,
- No undocumented steps exist,
- The applied logic matches documented specifications.

This enables reproducible reconstruction of model inputs for any historical decision.

Appendix B.3 Governance Control Enforcement (Gt)

Governance requirements are encoded as executable policy predicates that must be satisfied prior to model training or deployment.

Table 3: Policy Predicate Evaluation Log

Predicate	Threshold	Observed Value	Result
Missing Value Rate	≤ 2%	0.7%	Pass
Sensitive Attribute Balance	±5%	±3.1%	Pass
Label Distribution Shift	≤ 4%	1.9%	Pass

Audit Verification Test

The auditor confirms that dataset D_1''' was only released for model use after all predicates evaluated to "Pass," establishing that governance controls operated as enforceable gates rather than post-hoc documentation.

Appendix B.4 Monitoring and Drift Evidence (Mt)

Post-deployment, T-ETL continuously monitors statistical properties of production data against the baseline training distribution.

Table 4: Drift Monitoring Snapshot

Metric	Baseline (Training)	Current (Production)	Threshold	Status
Mean Feature Shift	—	1.6σ	≤ 2.0σ	Normal
Population Stability Index	—	0.12	≤ 0.2	Normal

Monitoring records are timestamped, versioned, and preserved, enabling auditors to verify that material distribution shifts are detected and escalated in a timely manner.

Appendix B.5 Provenance Tuple Assembly (Πt)

At any decision time t, the Evidence and Audit Layer allows reconstruction of:

$\Pi t = (D1, TDAG, Glogs, Mtelemetry)$

This enables an independent reviewer to verify:

- Which dataset was used,
- How it was transformed,
- What governance controls were enforced,
- Whether data behavior remained stable.

Appendix B.6 Assurance Implications

This worked example demonstrates how T-ETL converts abstract governance requirements into verifiable audit evidence. The framework does not depend on specific tooling or platforms; instead, it defines what must be provable for data governance to be regulatorily defensible. From an assurance perspective, the key outcome is that governance claims can be validated without reliance on narrative explanations or discretionary attestations, supporting independent audit, supervisory review, and forensic investigation.

Appendix C: Mapping Regulatory Requirements to T-ETL Evidence Artifacts

Table 5:

Regulation	Core Requirement	T-ETL Evidence Artifact
SR 11-7	Data suitability & control	D_t hash + lineage DAG
BCBS 239	End-to-end traceability	DAG traversal
EU AI Act	Bias & representativeness	Policy predicate logs
DORA	Data integrity & resilience	Immutable audit ledger

References

1. Board of Governors of the Federal Reserve System. SR 11-7: Guidance on Model Risk Management. Washington (DC): Federal Reserve; 2011.
2. Basel Committee on Banking Supervision. BCBS 239: Principles for Effective Risk Data Aggregation and Risk Reporting. Basel: Bank for International Settlements; 2013.
3. Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge (MA): MIT Press; 2016.
4. Gama J, Žliobaitė I, Bifet A, Pechenizkiy M, Bouchachia A. A survey on concept drift adaptation. ACM Computing Surveys. 2014;46(4):44.
5. European Commission. Regulation (EU) 2024/1689 (Artificial Intelligence Act). Official Journal of the European Union; 2024.
6. European Union. Regulation (EU) 2022/2554 on Digital Operational Resilience Act (DORA). Official Journal of the European Union; 2022.
7. Nasser M, *et al*. Big Data Governance: A systematic review of frameworks and maturity models. IEEE Access. 2025.
8. Zhou A, *et al*. Privacy-preserving verification of ML preprocessing via model-behavior indicators. IEEE Computer Society. 2025.

How to Cite This Article

Redu P. Auditing data governance for AI/ML in financial institutions: verifying the integrity, traceability, and lineage of training and production data under regulatory mandates. International Journal of Multidisciplinary Research and Growth Evaluation. 2026;7(3):42–56. doi:10.54660/.IJMRGE.2026.7.3.42-56.