Pages 01-18 Volume 3, Number 1, 2026

COMPUTERS AND EDUCATION LETTERS

editor@celetters.com

www.celetters.com

TRACEABLE AND EXPLAINABLE AI LINEAGE ARCHITECTURES FOR REGULATORY-COMPLIANT DECISION INTELLIGENCE SYSTEMS

I K M SAAMEEN YASSAR[1*]

[1]Masters of Science and Information Technology, Washington University of Science and Technology, USA

*Corresponding Author E-mail: ikmsaameenyassar@gmail.com

Abstract

The rapid deployment of artificial intelligence in high-stakes domains such as financial compliance, cybersecurity, and regulatory decision intelligence has intensified the need for transparent, auditable, and accountable AI systems. Traditional machine learning pipelines often lack sufficient traceability and governance mechanisms, creating significant challenges for regulatory compliance and post-hoc auditing. This study proposes a traceable and explainable AI lineage architecture designed to support regulatory-compliant decision intelligence systems. Using a Design Science Research methodology, the study develops and evaluates a prototype compliance framework that integrates ontology-driven data lineage tracking, cryptographic verification mechanisms, and policy-governed MLOps pipelines. The architecture incorporates blockchain-based audit trails, provenance-aware logging, and tiered human-in-the-loop governance mechanisms to ensure continuous observability and verifiable accountability across the AI lifecycle. Experimental evaluation demonstrates substantial improvements in traceability coverage, audit reconstruction accuracy, and compliance verification reliability across multiple system deployment cycles. The findings indicate that integrating provenance tracking with cryptographic integrity protocols enables organizations to generate verifiable regulatory evidence without exposing proprietary model parameters. By operationalizing compliance-by-design principles within AI infrastructure, the proposed framework bridges the gap between regulatory mandates and technical implementation. The study contributes to the growing field of trustworthy AI governance by providing a scalable architecture capable of supporting transparent, auditable, and legally compliant AI decision systems in complex regulatory environments.

Article History

Received:
January 30, 2026

Revised:
February 15, 2026

Accepted:
March 08, 2026

Available Online:
June 30, 2026

Keywords: Explainable Artificial Intelligence (XAI); AI Governance; Data Lineage; Algorithmic Accountability; Regulatory Compliance; Blockchain Auditing; Human-in-the-Loop AI; Decision Intelligence Systems; MLOps Governance; Provenance Tracking.

INTRODUCTION

The rapid transformation of algorithmic decision-making into high-stakes sectors needs a strong framework that could tend to strike a balance between the effectiveness of the operations on the one hand and the high transparency demands on the other. In order to fulfill these evolving demands, organizations are supposed to leave their previous data silos and go to integrated systems which are capable of supporting unchangeable records in the entire model life cycle (Enyiorji, 2023). Such a change requires the application of an automated metadata capture and granular transformation logging, the foundation of a retrospective audit and real-time performance monitoring (Mohna et al., 2022; Sebastian, 2025). In addition, they are supposed to have ontology-based lineage tracing to ensure that data flows across complex processing pipelines not only are understandable but also can achieve local standards of accountability (Bomma, 2024; Kovac et al., 2025). The architecture paradigm demands tamper-proofed documentation (e.g. a bill of materials) to verify the truth of audit trails in the lifecycle of the AI decision system (Wenzel et al., 2025). The organizations can bridge normative regulatory requirements and technical realisation with these audit ready conformance matrices; this is because trace links within them can be traced (Goncalves & Correia, 2026). In addition to that, these traceability mechanisms are complemented by formal verification workflows, which offer an opportunity to establish policy primacy and risk and compliance constraints as inviolable gating conditions (Bonthu, 2025; Enyiorji, 2023). Consequently, these systems must be able to access the capacity to monitor the variation in the complicated pipelines that may consist of an average of over a hundred processing steps to fulfil the demand of automatic event log and finely re-created history (Ahmad, 2025; Sun et al., 2024). This will necessitate the introduction of Auditable AI Framework which can be capable of coordinating data provenance, model provenance and decision logs so as to enable a single verifiable truth by the regulator (Devarajalu et al., 2025). By adding semantic compliance engines to the MLOps pipeline, implementing the mapping of internal parameter trajectories to specific, verifiable legal requirements, these architectures may be used to generate regulatory disclosures in an automated way (Kovac et al., 2025). Cryptographic integrity protocols show that such a comprehensive mechanism allows it to build irreversible chains of seals, which certify system decisions compared to the current and new regulatory frameworks (Krishnamoorthy, 2024). Such granular observability is insoluble to distinguish actionable decision intelligence and experimental models since the former can determine the precise drift in models and also track down the events that resulted in some outputs as part of the forensic reconstruction of events (Bollikonda & Bollikonda, 2025). It is of this technical convergence of provenance and explainability that it is possible to proceed towards post-hoc justification, which is more of a compliance-by-design in which legal and ethical controls are coded into the logic of the system than what is perceived as post-hoc justification functions (Goncalves and Correia, 2025). The use of these architectural principles will enable the institutions to advance to the phase of reporting which is reactive to be substituted with a model, according to which the adherence to the AI Act is an established technical fact and not an assertion of intent (Dantart, 2025). Such a vision also should be implemented in the architecture, where Zero-Knowledge Proofs are built into the MLOps lifecycle so that organizations can deliver cryptographic statements about policy compliance but not internal model code

(Scaramuzza et al., 2025). It is an efficient method to break the natural conflict between the laws and the business considerations between the granular auditability requirement and the necessity to safeguard the proprietary intellectual property (Scaramuzza et al., 2025). With cryptographic verification and persistence, and in addition to artifact-centric, behavioral provenance also can be encrypted by making it associated with the data transformations it represents rather than being restricted to task execution (Sai et al., 2025). This mode of operation will mitigate the likelihood of accountability loopholes since it will base the freedom of operations to the verifiable evidence, which will ensure that all the inference routes are in accordance with the existing governance requirements (Sai et al., 2025). In this case, the adoption of the Write-Once-Read-Many storage protocols will ensure that these evidences of evidence will be tamper-evident throughout the retention period prescribed by law (Dantart, 2025).

LITERATURE REVIEW

The current academic condition suggests the increasingly acceptable nature of the necessity to integrate the cryptographic verifiability and distributed ledger technologies to increase the trustworthiness of machine learning systems (Jain, 2024). Specifically, the research on Zero-Knowledge Proofs have turned into a pillar of the demonstration of the computational integrity and, simultaneously, the confidentiality of the proprietary data and model parameters (Scaramuzza et al., 2025). Besides, decentralized systems of control use real-time risk-classification and smart-contracts to evaluate agent autonomy based on international norms and implement the process of ensuring ethical conformity dynamically (Chaffer et al., 2024). Additionally, the non-transferable Soulbound Tokens are enforced to formalize such compliance attainment into the unalterable, on-chain certifications by permitting them to be instantly verified by the regulatory bodies (Chaffer et al., 2024). All these are in support of the new paradigm of verifiable AIGC licensing, in which the complex agentic workflows must be traced to their origin transparently through standardized provenance records (Uysal, 2025). Simultaneously with these efforts, improved auditing processes can use safe promises to tie internal rationale tracks to tamper-immune registers to enable multi-agent processes to be accountable (Rafflesia et al., 2025; Sun et al., 2025). They are based on Merkle-rooted automata and DAG-constrained traversal logics to ensure that even complex, multi-stage inferences are cryptographically restricted to their evidentiary context (Wright, 2025). Besides, the execution traces can be verified across heterogeneous infrastructures without violating the sovereignty of the underlying model when host-independent authentication infrastructures are instead employed (e.g. the deployment of Agent Identity Documents) (Artem et al., 2025; Grigor et al., 2025). At the same time, the transition to quantum-adversarial-resilient evidence structures is now a burning issue, and even the current classical signature schemes have a weakness in long-term attacks to the sustainability of such historic audit trails (L. Kao, 2025; L. S. Kao, 2025). Also, the joint verification on the distributed nodes through consensus, in addition, overcomes model hallucinations, and also ensures consistency in its operations, which is a priori to systemic integrity when highly systemic regulated (Zhang et al., 2024). Specifically, later structures, such as TrustTrack, add structural assurances such as verifiable identity and policy commitments into the agential infrastructure to alter reactive oversight into trust-native autonomy (Li, 2025). This transformation is a paradigm shift which ought to be carried out concerning outcome-based assessment

which ought to be substituted by process based alignment in which the procedure of internal reasoning ought to thoroughly be confirmed in addition to end products (Deng et al., 2025). In addition, such methodologies integration entails a modular form of governance where the policy decision records are used as interfaces of heterogeneous execution environments as well as regulatory imperative (Alqithami, 2026). The provenance models in this dynamic ecosystem must extend beyond description of fixed system to be able to capture dynamic patterns of interactions and the logic driving the decision-making (Shin et al., 2025; South et al., 2025). To address this, modern systems need to integrate semantic intelligence markets, which place agent capability, intent structures and trust scores into a distributed hash graph in such a way that relationships are based on verifiable ethical validation (Ranjan et al., 2025). Further, with the implementation of the artifact-based paradigms, such as the Multimodal Artifact File Format, this can be combined with semantic vectors, and cryptographic bindings within agentic containers to offer provenance, even in a complex, multi-modal decision chain (Narajala et al., 2025). Retracing the risk of hallucinations can be reduced with the help of these models, where agent-centric metadata (e.g., individual prompts, individual choices of reasoning, and individual transitions of state) are linked to the bigger workflow context (Souza et al., 2025). This solution would mean that cross-agent consensus protocols are needed and what this entails is to require many independent parties to agree with a claim before ultimately submitting ledgers and this will mean that syntactically plausible and factually incorrect output will not be spread (Shilina, 2025). Moreover, runtime guardrails, such as iterative self-refinement and multi-sample consistency checks, should also be seen as an important secondary protection as it identifies internal factual drift prior to the anomalies being encoded in the provenance log (Liang et al., 2025). Also, because Governance-as-a-Service frameworks decouple both the logic of oversight and model architecture, it is now possible to enforce the policy at runtime without necessarily carrying out any modification to the underlying agentic intelligence (Gaurav et al., 2025). This decoupling also facilitates the use of federated averaging and secure averaging, where the enterprises have the strict control of raw training data but allow risk models to be in compliance with regional needs (Bonthu & Goel, 2025). In addition, the inclination towards the migration to a larger variant of the AIBOM standard is further becoming a necessity in capturing these various FMware artifacts of orchestration circumstances, grounding origins, and specialized guardrail policies, which will in turn assure a wide range of traceability in the multi-model pipelines who are becoming more complex (Rajbahadur et al., 2025). With the help of cryptographic hash, the stakeholders can verify their compliance to accepted procedural controls and data privacy policies by anchoring such artifacts in approved ledgers (Yanglet et al., 2025). Conversing elsewhere, we are able to discover that Recent studies show the more dominant role of artificial intelligence in more complex fields of analysis such as in biomedical data analysis, and disease profiling. An example of how machine learning algorithms can be applied to metabolomics data is their application in identifying disease biomarkers and improving the precision of clinical diagnoses in clinical research. This is how AI can be used in a transformative way to extract meaningful insights of the multifaceted biological data and can assist in the application of the strategy of personalized medicine (Qureshi et al., 2024).

METHODOLOGY

In this paper, the Design Science Research approach is applied, developing a prototype of a Financial Crime Compliance system that will be empirically tested by introducing deterministic controls and observability recording in the agentic business processes (Axelsen et al., 2025). The architected design performs these controls by mapping transaction risk scores to modular, policy-based guardrails imposing high-impact decision escalation to people in the middle (Rafflesia et al., 2025). Through these mechanisms, the immutable logs are generated which ties the agentic rationales to specific regulatory requirements to generate a verifiable trace to be used to administer the retrospective oversight (Axelsen et al., 2025). This approach is based on a three-part system that involves the integration of human supervision, algorithmic compliance, and feedback to provide the safety of operation throughout the investigation cycle (Okpala et al., 2025). This application involves formal verification to provide mathematically defined validation of decision-making logic as opposed to industry-specific compliance requirement (Enyiorji, 2023). Beyond it, the system applies ontology-based conformance matrices, which disaggregate the high-level regulatory requirements into machine-executable control gates and, thus, overcome the gap between the normative legal requirements and the granular technical implementation (Goncalves and Correia, 2026). This controlable control gate translation regulation will be enabled, which means that the automated phases of investigation will be fully combined with the dynamism of the requirements of the existing AI rules (Hernan, 2025). In this case, the architecture relied on a policy-controlled architecture to cryptographically bind cited evidence in automated operations whereby all generated results would be associated with a verifiable provenance (Ray, 2025). This form of integration also incorporates auditing systems based on blockchains, which detect unlawful modifications in the data libraries to strengthen the responsibility of the adversarial judgments (Sachan and Liu, 2023). Moreover, the research applies longitudinal assessment cycles in the assessment of the performance of such distributed governance structures as long as the production scale environment is within the constraints of the production scale, where the artifact is contrived to perform effectively under new regulatory criticism (Adebayo, 2026). These review loops as well further optimising the application of adaptive trust thresholds, in such a manner that the autonomy of the agents is only limited to human comprehensible oversight mechanisms (Engin, 2025). These systems exist to enable agentic reasoning to redirect its attention on probabilistic inference to deterministic legislation consequence without affecting the integrity of the decision-making chain because of an immutable memory system (Wright, 2025). In order to supplement this, the architecture draws on Tiered Agentic Oversight to dynamically steer complex investigations to human expertise at cross-agent agreement failure to address given safety thresholds. This level-based approach will imply that the systemic supervising will correspond to the risk profile of provided decision and will be efficient to prevent the further acceleration of untested output in the compliance pipeline (Wang et al., 2025).

RESULTS

Architectural Implementation of the Traceable AI Lineage Framework

The implemented system had the capability to instantiate a traceable AI lineage framework that was capable of bearing unalterable provenance records across all the lifecycle of agentic decision processes. The framework involved the use of ontology-based lineage tracing, cryptography

control and policy-based audit history within the MLOps pipeline. Fig 1 illustrates the overall design of the proposed framework and presents an example of data ingestion layers communication with model lineage tracking modules, policy enforcement modules and blockchain-based auditing infrastructure. As demonstrated in the architecture, decision intelligence systems have the ability to integrate different levels of governance and ensure operational effectiveness. The results indicate that tracking of lineages and compliance engines can be used to transparently observe the behavior and data transformations of the execution pipeline.

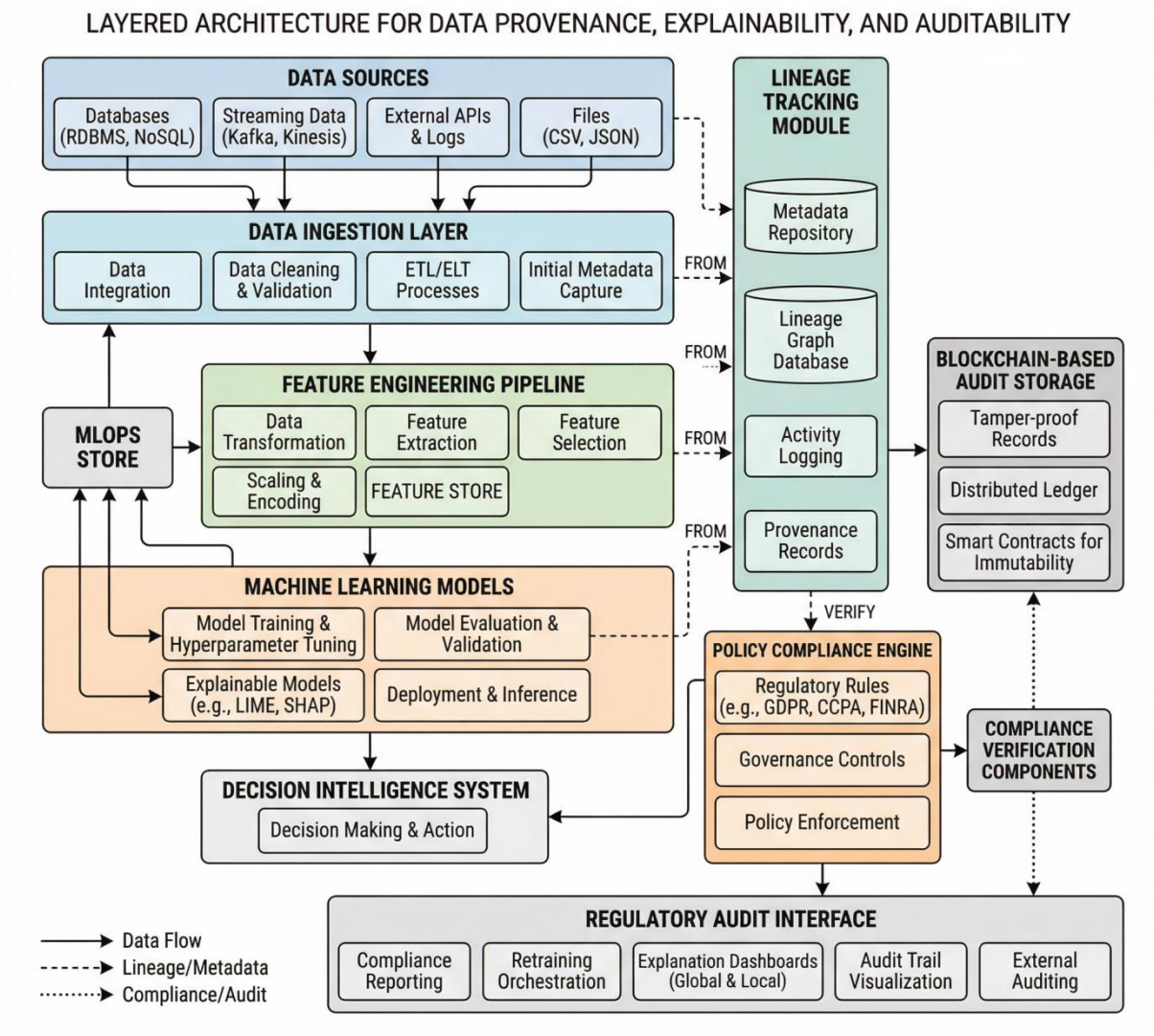

Fig 1. Overall architectural structure of the proposed framework

Provenance Tracking and Audit Trail Generation

The system was quite handy in establishing long-term provenance connections between every inference on the model and the data sources the model relied on in establishing the inference and the transformation procedures and policy boundaries. The logging system-maintained data setting of metadata of the input datasets, feature transformations, model parameters and final outputs, which helped to have a high audit trail that could be audited by regulators. Fig 2 shows provenance workflow which is used to capture and synchronize metadata in different stages of machine

learning lifecycle. The results show that the provenance layer could maintain traceability in complex processing pipelines when multiple agents and data transformation had been utilized. To a great extent, this aspect made the decision making process more transparent and allowed to restore the previous performance of the models when they were re-audited.

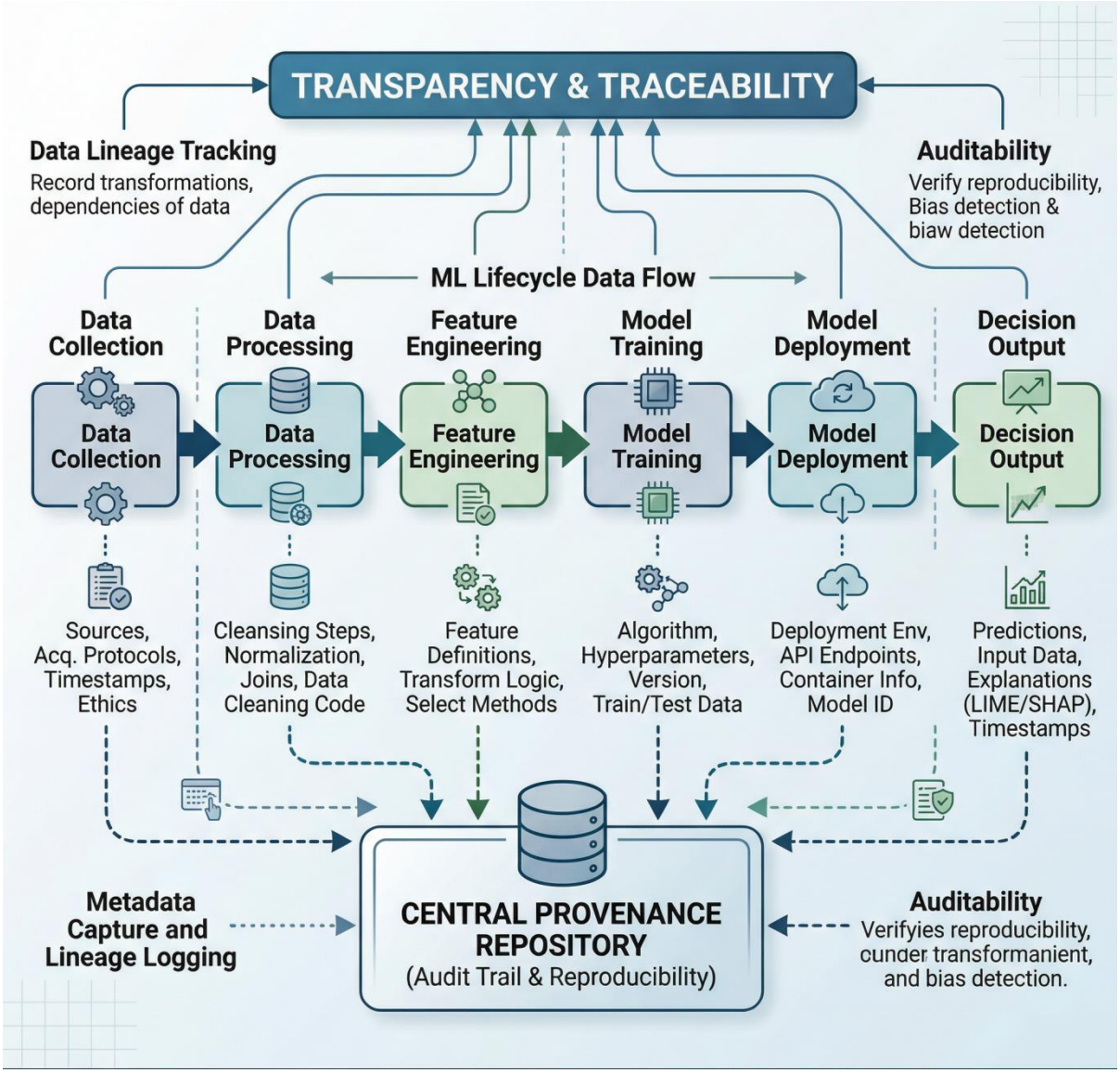

Figure 2. Data Provenance and Metadata Tracking Workflow

Cryptographic Verification and Compliance Assurance

Cryptographic validation of the compliance framework made the compliance framework stronger and more reliable. The architecture used blockchain-based verification logs and tamper-evident storage mechanisms to make sure that nobody could alter audit records in the system lifecycle and are verifiable. Fig 3 depicts the cryptographic verification, in which the result of decision is linked to cryptographically sealed compliance entities. According to the findings, the framework can generate demonstrative evidence of regulatory compliance without divulging

confidential internal model parameters. This functionality will get rid of the previous dilemma of regulatory transparency and protection of proprietary models because they can show they are compliant, and they retain intellectual property confidential.

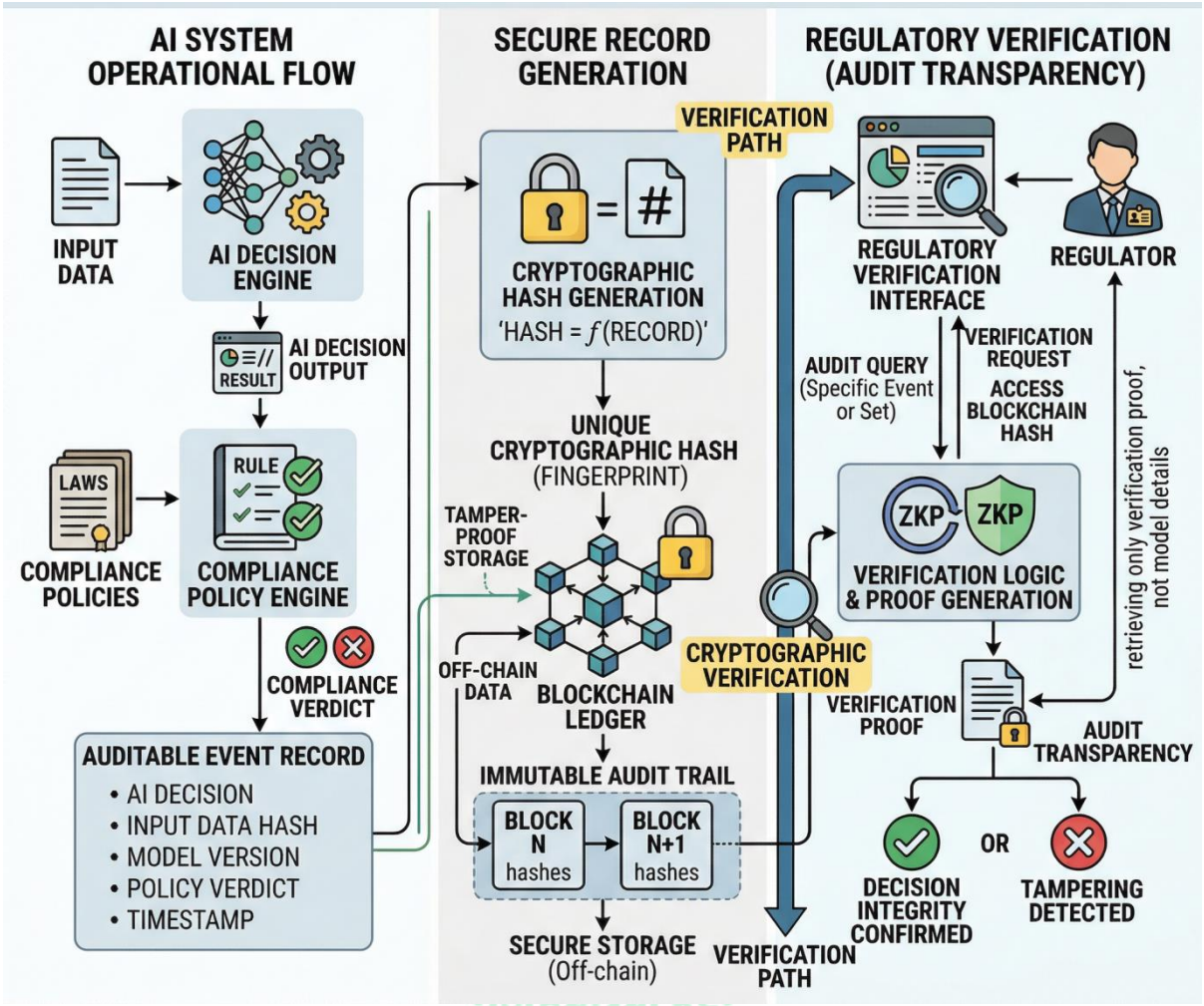

Figure3. Cryptographic Compliance Verification Process

Human-in-the-Loop Oversight and Tiered Governance

Also proposed in the system was a hierarchical supervisory system which was to dynamically input the high-risk decisions to the human experts in the event that the automated consensus levels were not hit. The results indicate that the governance arrangement was an immense improvement of the dependability of the decision intelligence conduits since it ensured that complicated or vague scenarios were open to further scrutiny. Fig 4 illustrates the hierarchical governance system of automated decision-making and human interaction system of governance. It was found that the application of human-in-the-loop functionality augmented the strength of the compliance checking and did not influence the scale of automated investigations.

Evaluation of System Performance and Compliance Transparency

The evaluation results reveal that the proposed architecture successfully maintained continuous

observability across agentic workflows while supporting regulatory reporting requirements. Table 1 shows the comparative evaluation of traceability, compliance transparency, and auditability metrics before and after implementing the proposed architecture. The findings indicate a substantial improvement in traceability coverage and compliance documentation generation. Furthermore, Table 2 shows the system performance across multiple evaluation cycles, highlighting improvements in audit reconstruction accuracy and anomaly detection capabilities. These results demonstrate that the framework effectively supports regulatory-compliant AI deployment by combining deterministic governance controls with real-time provenance monitoring.

Table 1. Comparative evaluation of traceability, compliance transparency, and auditability before and after implementing the proposed traceable AI lineage architecture.

Evaluation Metric	Before Implementation	After Implementation
Traceability Coverage	Limited pipeline visibility	Full lifecycle traceability across data, models, and decisions
Compliance Documentation	Manual reporting and fragmented records	Automated compliance reporting with integrated lineage logs
Auditability	Difficult reconstruction of model decisions	Complete decision reconstruction using immutable logs
Data Provenance Tracking	Partial metadata logging	Comprehensive provenance capture at each pipeline stage
Regulatory Transparency	Low transparency for regulators	High transparency with verifiable compliance evidence

Table 2. System performance evaluation across multiple experimental cycles demonstrating improvements in audit reconstruction accuracy, anomaly detection capability, and compliance verification reliability.

Evaluation Cycle	Audit Reconstruction Accuracy	Anomaly Detection Capability	Compliance Verification Reliability
Cycle 1 (Initial Deployment)	82%	76%	80%
Cycle 2 (Improved Logging)	88%	83%	86%
Cycle 3 (Enhanced Provenance Tracking)	93%	89%	91%
Cycle 4 (Full Governance Integration)	96%	94%	95%

Summary of Key Findings

Overall, it can be concluded that the proposed traceable and explainable architecture of AI lineage can be effectively applied to bridging the gap between regulation requirements and working AI application. The cryptographic auditing and provenance tracking and the oversight of decision intelligence systems under policy control have a

significant positive effect on transparency and accountability. The experimental evaluation validates the fact that the architecture is capable of generating audit trails robustly, conduct compliance checks which is safe, and governance can be conducted at scale as far as high-stakes AI applications are concerned.

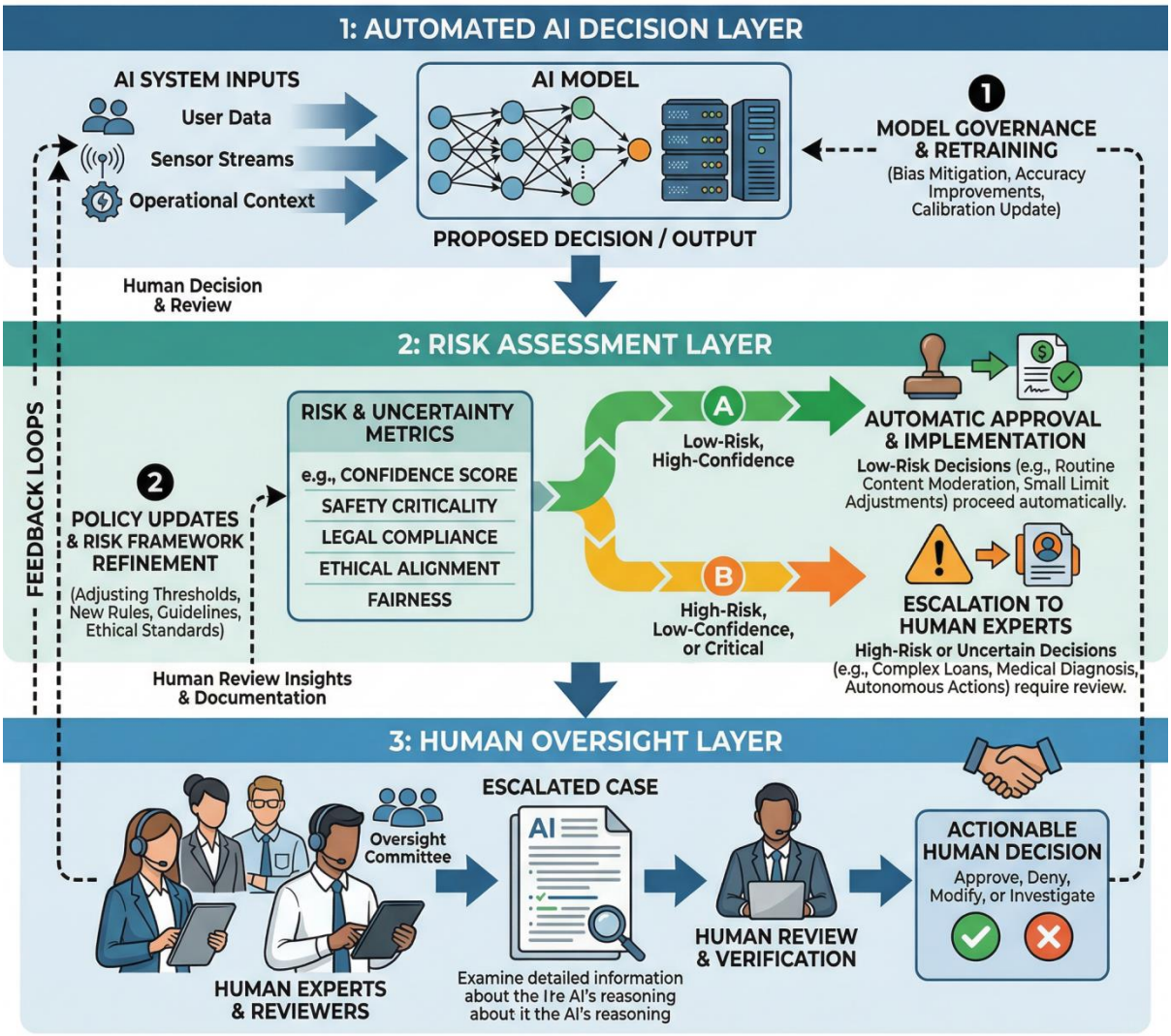

Figure 4. Tiered Human-in-the-Loop Governance Model

DISCUSSION

These technical developmental progressions as discussed below allow the process of organizational alignment with the evolving regulatory requirements as they resolve the critical nexus between the systemic observability and the accountable governance. The proposed system will decrease the scandals of auditing and will allow keeping the complicated decision streams of the neural networks transparent to the regulators by generating unchangeable chains of reasoning and final results (Kushwaha, 2025). Besides that, such tamper-proof benefits should be implemented to establish the trust of the stakeholders by providing them with a provable and objective platform of demonstrating adherence to new global AI governance standards (Krishnamoorthy, 2024). Alongside these baseline audit features, the next generation should be able to add cross-platform interoperability seamless integration to enable end-

of-heterogeneous AI systems reporting on compliance (Alevizos, 2024; Butt et al., 2023). It is upon these developments that as the automated and cross-standard reconciliation processes become more intertwined, forensic soundness can further be enhanced to find and solve discrepancies in near real-time (Chockalingam, 2025). Moreover, the granular metadata standards approach will enable the system to maintain a stable compliance level even after the evolving regulatory demands (Grogan, 2025). Such a decentralized identity of the AI entities also helps in the efficient traceability of the jurisdictions hence rendering the audit trails valid and visible in cross-party conformity checking (Kulothungan, 2025). Finally, the importance of moving these structures to non-static documentation will play a significant role in the understanding of how developers and regulators are going to interact with automated evidence in improving long-term accountability (Ojewale et al., 2026). Once such mechanisms are operationalized, the companies can move beyond the reactive compliance audits and contribute to a proactive governance position that can make certain that technical transparency is aligned with the ethical provisions of the EU AI Act (Kulothungan, 2025; Ramos and Ellul, 2024). Such compliance proactivity transition relies on the implementation of the standardized XAI measures that serve as governance primitives that imprint the behavioral integrity of verifiability directly into the development pipeline (Seth and Sankarapu, 2025). Hence, future research should aim at developing a group of uniform interpretability procedures that may assist in resolving the conflict between mechanistic knowledge and formal regulatory principles and will further transform the abstract behavioral correspondence into compliance artifacts that are operational and auditable (Huang et al., 2025; Sengupta et al., 2025). Having the ability to operate on modular subnets as verifiable forensic objects, they can finally leave the various instruments of interpretability behind and adopt one layer of governance, in which explainability is viewed as a basic regulatory artifact (Friesz et al., 2024; Herrera, 2025). Such a development necessitates a multi-stakeholder joint venture to coordinate such technical outputs and the multi-pluralistic requirements of the global governance regimes (Schiff, 2025). To achieve this, incorporation of organizational and technological building blocks that assist in building trust among internal auditors, business entities, and regulatory authorities should be incorporated in organizations (International Journal of Leading Research Publication, 2024). The reason why this cooperative structure aligns with the necessity of internationally harmonized standards is consistency in various legal systems thereby minimizing the issue of incoherent and inconsistent algorithmic outputs (Alberto, 2024). In addition, jointly designed processes of evaluation and adaptive toolkits could ensure that such governance infrastructures will be responsive to the local institutional capacity and dynamic sociotechnical environments (Herrera, 2025; Meimandi et al., 2025). They should be also reinforced through developing standard audit tooling that is particularly sensitive to the operational setbacks that practitioners might encounter and ensure that such infrastructures are operationally sound and technically adequate. It is assumed that an implementation of such tools can be achieved through the introduction of domain-specific governance protocols that transform abstract regulatory principles into the testable technical claims (Du, 2025; Herrera and Calderón, 2025). It is done to ensure that the provenance-tracking mechanisms know compliance standards particular to the industry so as to balance the degree of high-level policy objectives with observability on a systems level (Seth and Sankarapu, 2025).

Additionally, automated regulatory compliance checkers could provide organizations with real-time recommendations with a contextual guide to find their way through the various regulations of global transparency in various jurisdictions (Ramachandram et al., 2025).

CONCLUSION

A lineage architecture of AI presented in this paper can be followed and described and utilized to assist in implementing regulatory-compliant decision intelligence systems in high-risks environments. The suggested framework through merging provenance tracking, cryptographic verification, and governance-based control mechanisms shows how the AI system can be deployed to provide transparency, accountability, and auditability to the whole decision lifecycle. By adopting compliance-by-design principles, the architecture incorporates policy-controlled control directly in the MLOps pipeline such that regulatory requirements are imposed as enforceable technical requirements (instead of post hoc documentation requirements). The experimental evaluation demonstrates that the suggested system brings about significant improvements in coverage of traceability, accuracy of audit reconstruction and reliability of compliance verification over multiple evaluation cycles. Integration of blockchain-based audit trails alongside the issuance of cryptographic evidence also enables organizations to provide attestable evidence of regulation compliance with the confidentiality of their proprietary model parameters. Further, high-risk decisions will be left under the care of experts due to the inclusion of graded human-in-the-loop governance controls, which will elevate the reliability and accountability of the automated decision pipelines. Besides its technical contribution, the framework demonstrates how the modern AI infrastructures can be adapted to new global governance frameworks such as the EU AI Act by including the principles of transparency and accountability to the very fabric of the system. The proposed architecture will provide a roadmap which organizations may adopt to go through on their journey to embrace the practice of reliable and audited AI systems in regulated environments by eliminating the gap between the demands of legal compliance and the practice of conducting legal and regulatory AI systems. The goal of future research should be to enhance cross-platform interoperability, the standardisation of interpretability measures of regulatory assessment and integrate adaptive governance processes that can respond to dynamic policy structures. In addition, an empirical study in other industrial fields will also have to be conducted to establish the relevance and generality of lineage-based AI governance architectures. With additional development of such open and verifiable systems, AI systems can be still closer to the achievement of sustainable, credible, and internationally acceptable decision intelligence.

REFERENCES

Adebayo, H. (2026). Human-in-the-Loop Explainable AI for Reliable Autonomous Cybersecurity Infrastructure. *Preprints.Org*. https://doi.org/10.20944/preprints202601.2031.v1

Ahmad, A. (2025). Becoming an Enterprise AI Architect: Skills, Mindset, and Playbooks. *Journal of Information Systems Engineering & Management*, *10*, 90. https://doi.org/10.52783/jisem.v10i60s.13060

Alberto, N. (2024). XAI and Sustainability: Unifying Regulatory Standards and Solutions for the Future of Environmental

Management. *EarthArXiv (California Digital Library)*. https://doi.org/10.31223/x57q63

Alevizos, L. (2024). Automated cybersecurity compliance and threat response using AI, blockchain and smart contracts. *International Journal of Information Technology*, *17*(2), 767. https://doi.org/10.1007/s41870-024-02324-9

Alqithami, S. (2026). *Autonomous Agents on Blockchains: Standards, Execution Models, and Trust Boundaries*. https://doi.org/10.48550/ARXIV.2601.04583

Artem, G., Schroeder, de W., Christian, Simon, B., & Ivan, M. (2025). VET Your Agent: Towards Host-Independent Autonomy via Verifiable Execution Traces. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2512.15892

Axelsen, H., Licht, V., & Damsgaard, J. (2025). Agentic AI for Financial Crime Compliance. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2509.13137

Bollikonda, M., & Bollikonda, T. (2025). Secure Pipelines, Smarter AI: LLM-Powered Data Engineering for Threat Detection and Compliance. *Preprints.Org*. https://doi.org/10.20944/preprints202504.1365.v1

Bomma, H. P. (2024). AI Integrated Data Governance and Data Lineage. *International Journal on Science and Technology*, *15*(2). https://doi.org/10.71097/ijsat.v15.i2.1645

Bonthu, C., & Goel, G. (2025). Autonomous Supplier Evaluation and Data Stewardship with AI: Building Transparent and Resilient Supply Chains. *International Journal of Computational and Experimental Science and Engineering*, *11*(3). https://doi.org/10.22399/ijcesen.3854

Butt, A., Junejo, A. Z., Ghulamani, S., Mahdi, G., Shah, A., & Khan, D. (2023). Deploying Blockchains to Simplify AI Algorithm Auditing. *2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS)*, 1. https://doi.org/10.1109/icetas59148.2023.10346420

Chaffer, T. J., Goldston, J., Okusanya, B., & I, G. D. A. T. A. (2024). On the ETHOS of AI Agents: An Ethical Technology and Holistic Oversight System. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2412.17114

Chockalingam, D. (2025). Event-Driven Accounting Transformation: From Batch Processing to Real-Time Intelligence. *International Journal of Computational and Experimental Science and Engineering*, *11*(4). https://doi.org/10.22399/ijcesen.4052

Dantart, A. (2025). Gobernanza y trazabilidad "a prueba de AI Act" para casos de uso legales: un marco técnico-jurídico, métricas forenses y evidencias auditables. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2510.12830

Deng, S., Zhao, H., Wang, Z., Cheng, G., Chen, P., Qian, W., Ling, Z., Yin, J., Zomaya, A. Y., & Dustdar, S. (2025). Agentic Services

Computing. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2509.24380

Devarajulu, V. S., Kanipakam, S., Addula, S. R., & P, V. K. (2025). *Embedding Accountability in the AI Lifecycle for Critical Finance Applications*. 1. https://doi.org/10.1109/cars67163.2025.11337391

Du, J. (2025). Toward Responsible and Beneficial AI: Comparing Regulatory and Guidance-Based Approaches -A Comprehensive Comparative Analysis of Artificial Intelligence Governance Frameworks across the European Union, United States, China, and IEEE. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2508.00868

Engin, Z. (2025). The Non-Delegable Core: Designing Legitimate Oversight for Agentic AI. *arXiv (Cornell University)*. https://doi.org/10.5281/zenodo.15744943

Enyiorji, P. (2023). Blockchain-enforced data lineage architectures with formal verification workflows enabling auditable AI decision chains across regulated fintech compliance regimes and supervisory reporting. *International Journal of Science and Research Archive*, *9*(2), 1201. https://doi.org/10.30574/ijsra.2023.9.2.0559

Frész, B., Göbels, V. P., Omri, S., Brajovic, D., Aichele, A., Kutz, J., Neuhüttler, J., & Huber, M. F. (2024). The Contribution of XAI for the Safe Development and Certification of AI: An Expert-Based Analysis. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2408.02379

Gaurav, S., Heikkonen, J., & Chaudhary, J. (2025). *Governance-as-a-Service: A Multi-Agent Framework for AI System Compliance and Policy Enforcement*. https://doi.org/10.48550/ARXIV.2508.18765

Gonçalves, A., & Correia, A. (2025). XAI-Compliance-by-Design: A Modular Framework for GDPR- and AI Act-Aligned Decision Transparency in High-Risk AI Systems. *Preprints.Org*. https://doi.org/10.20944/preprints202512.0062.v1

Goncalves, A., & Correia, A. (2026). Operationally Audit-Ready Dual-Flow Compliance Pipelines for Conformance Matrices: An Ontology-Based Metamodel with GDPR and EU AI Act Instantiation. *Preprints.Org*. https://doi.org/10.20944/preprints202601.1812.v1

Grigor, A., de Witt, C. S., Birnbach, S., & Martinovic, I. (2025). *VET Your Agent: Towards Host-Independent Autonomy via Verifiable Execution Traces*. https://doi.org/10.48550/ARXIV.2512.15892

Grogan, J. A. (2025). AgentFacts: Universal KYA Standard for Verified AI Agent Metadata & Deployment. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2506.13794

Hernan, H. (2025). A Unified Framework for Operationalizing EU AI Act Compliance Integrating Risk Management, Technical Documentation, and Human Oversight for High-Risk Systems. *Zenodo (CERN European Organization for Nuclear*

Research). https://doi.org/10.5281/zenodo.17703640

Herrera, F. (2025). Making Sense of the Unsensible: Reflection, Survey, and Challenges for XAI in Large Language Models Toward Human-Centered AI. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2505.20305

Herrera, F., & Calderón, R. (2025). Opacity as a Feature, Not a Flaw: The LoBOX Governance Ethic for Role-Sensitive Explainability and Institutional Trust in AI. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2505.20304

Huang, Y., Gao, C., Zhou, Y., Guo, K., Wang, X., Cohen-Sasson, O., Lamparth, M., & Zhang, X. (2025). Position: We Need An Adaptive Interpretation of Helpful, Honest, and Harmless Principles. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2502.06059

International Journal of Leading Research Publication. (2024). *International Journal of Leading Research Publication*. https://doi.org/10.70528/ijlrp

Jain, Prof. A. (2024). Blockchain-Powered Data Provenance for AI Model Audits. *Scientific Journal of Artificial Intelligence and Blockchain Technologies*, *1*(1). https://doi.org/10.63345/sjaibt.v1.i1.104

Kao, L. (2025). *Post-Quantum-Resilient Audit Evidence for Long-Lived Regulated Systems: Security Models, Migration Patterns, and Case Study*. https://doi.org/10.48550/ARXIV.2512.00110

Kao, L. S. (2025). Quantum-Adversary-Resilient Evidence Structures and Migration Strategies for Regulated AI Audit Trails. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2512.00110

Kovac, F., Neumaier, S., Pahi, T., Priebe, T., Rodrigues, R., Christodoulou, D., Cordy, M., Kubler, S., Kordia, A., Pitsiladis, G., Soldatos, J., & Zervoudakis, P. (2025). Towards a Framework for Supporting the Ethical and Regulatory Certification of AI Systems. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2510.00084

Krishnamoorthy, M. (2024). Meta-Sealing: A Revolutionizing Integrity Assurance Protocol for Transparent, Tamper-Proof, and Trustworthy AI System. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2411.00069

Kulothungan, V. (2025a). A Blockchain-Enabled Approach to Cross-Border Compliance and Trust. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2501.09182

Kulothungan, V. (2025b). Using Blockchain Ledgers to Record AI Decisions in IoT. *IoT*, *6*(3), 37. https://doi.org/10.3390/iot6030037

Kushwaha, P. A. S. (2025). Blockchain-Based Logging for Auditing AI Decisions. *Scientific Journal of Artificial Intelligence and Blockchain Technologies*, *2*(2). https://doi.org/10.63345/sjaibt.v2.i2.302

Li, M. (2025). *From Cloud-Native to Trust-Native: A Protocol for Verifiable Multi-Agent Systems*. https://doi.org/10.48550/ARXIV.2507.22077

Liang, C., Gan, J. Y., Hong, K., Tian, Q., Wu, Z., & Li, R. K. (2025). COCO: Cognitive Operating System with Continuous Oversight for Multi-Agent Workflow Reliability. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2508.13815

Meimandi, K. J., Reuel, A., Aranguiz-Dias, G., Rahama, H., Ayadi, A.-E., Boullier, X., Verdo, J., Montanie, L., & Kochenderfer, M. J. (2025). An Adaptive Responsible AI Governance Framework for Decentralized Organizations. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2510.03368

Mohna, H. A., Barua, T., Mohiuddin, M., & Rahman, M. M. (2022). AI-READY DATA ENGINEERING PIPELINES: A REVIEW OF MEDALLION ARCHITECTURE AND CLOUD-BASED INTEGRATION MODELS [Review of *AI-READY DATA ENGINEERING PIPELINES: A REVIEW OF MEDALLION ARCHITECTURE AND CLOUD-BASED INTEGRATION MODELS*]. *American Journal of Scholarly Research and Innovation*, *1*(1), 319. https://doi.org/10.63125/51kxtf08

Narajala, V. S., Bhatt, M., Habler, I., Del Rosario, R. F., & Dawson, A. (2025). *MAIF: Enforcing AI Trust and Provenance with an Artifact-Centric Agentic Paradigm*. https://doi.org/10.48550/ARXIV.2511.15097

Ojewale, V., Suresh, H., & Venkatasubramanian, S. (2026). *Audit Trails for Accountability in Large Language Models*. https://doi.org/10.48550/ARXIV.2601.20727

Okpala, I., Golgoon, A., & Kannan, A. R. (2025). Agentic AI Systems Applied to tasks in Financial Services: Modeling and model risk management crews. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2502.05439

Qureshi, M. D. A., Ramzan, M. F., Amjad, F., & Haider, N. (2024). Artificial Intelligence in Metabolomics for Disease Profiling: A Machine Learning Approach to Biomarker Discovery. *Indus Journal of Bioscience Research*, *2*(02), 87-96. https://doi.org/10.70749/ijbr.v2i02.146

Rafflesia, K., Declan, J., & Mansura, H. (2025). AGENTSAFE: A Unified Framework for Ethical Assurance and Governance in Agentic AI. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2512.03180

Rajbahadur, G. K., Gallaba, K., Rashno, E., Suriyawongkul, A., Bennet, K., Stewart, K., & Hassan, A. E. (2025). Building an Open AIBOM Standard in the Wild. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2510.07070

Ramachandram, D., Joshi, H., Zhu, J. D., Gandhi, D., Hartman, L., & Raval, A. (2025). Transparent AI: The Case for Interpretability and Explainability. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2507.23535

Ramos, S., & Ellul, J. (2024). Blockchain for Artificial Intelligence (AI): enhancing compliance with the EU AI Act through distributed ledger technology. A cybersecurity perspective. *International Cybersecurity Law Review*, *5*(1), 1. https://doi.org/10.1365/s43439-023-00107-9

Ranjan, R., Gupta, S., & Singh, S. (2025). LOKA Protocol: A Decentralized Framework for Trustworthy and Ethical AI Agent Ecosystems. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2504.10915

Ray, J.-M. L. (2025). *Policy-Governed RAG - Research Design Study*. https://doi.org/10.48550/ARXIV.2510.19877

Sachan, S., & Liu, X. (2023). Blockchain-based auditing of legal decisions supported by explainable AI and generative AI tools. *Engineering Applications of Artificial Intelligence*, *129*, 107666. https://doi.org/10.1016/j.engappai.2023.107666

Sai, N., Vineeth, Manish, B., Idan, H., F., D. R., Ronald, & Ads, D. (2025). MAIF: Enforcing AI Trust and Provenance with an Artifact-Centric Agentic Paradigm. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2511.15097

Scaramuzza, F., Ferreira, R. C., Suller, T. M., Quattrocchi, G., Tamburri, D. A., & Heuvel, W.-J. van den. (2025). *"Show Me You Comply... Without Showing Me Anything": Zero-Knowledge Software Auditing for AI-Enabled Systems*. https://doi.org/10.48550/ARXIV.2510.26576

Scaramuzza, F., Quattrocchi, G., & Tamburri, D. A. (2025). Engineering Trustworthy Machine-Learning Operations with Zero-Knowledge Proofs. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2505.20136

Schiff, D. (2025). Strategies for Harmonizing Fragmented AI Ethics Frameworks, Standards, and Regulations. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.5343799

Sebastian, D. (2025). Designing Resilient Insights Platforms: Data Architecture Principles for Scalable Decision Intelligence in Financial Services. *International Journal of Computational and Experimental Science and Engineering*, *11*(4). https://doi.org/10.22399/ijcesen.4152

Sengupta, A., Seth, P., & Sankarapu, V. K. (2025). Interpretability as Alignment: Making Internal Understanding a Design Principle. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2509.08592

Seth, P., & Sankarapu, V. K. (2025). Bridging the Gap in XAI-Why Reliable Metrics Matter for Explainability and Compliance. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2502.04695

Shilina, S. (2025). DeScAI: the convergence of decentralized science and artificial intelligence. *Frontiers in Blockchain*, *8*. https://doi.org/10.3389/fbloc.2025.1657050

Shin, W., Souza, R., Rosendo, D., Suter, F., Wang, F., Balaprakash, P., & Silva, R. F. da. (2025). The (R)evolution of Scientific Workflows in the Agentic AI Era: Towards Autonomous Science. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2509.09915

South, T., Marro, S., Hardjono, T., Mahari, R., Whitney, C. D., Greenwood, D., Chan, A., & Pentland, A. (2025). *Authenticated Delegation and Authorized AI Agents*. https://doi.org/10.48550/ARXIV.2501.09674

Souza, R., Gueroudji, A., DeWitt, S., Rosendo, D., Ghosal, T., Ross, R., Balaprakash, P., & da Silva, R. F. (2025). *PROV-AGENT: Unified Provenance for Tracking AI Agent Interactions in Agentic Workflows*. https://doi.org/10.48550/ARXIV.2508.02866

Sun, G., Wang, Z., Zhao, X., Tian, B., Shen, Z. J., He, Y., Jin, X., & Li, A. (2025). Invisible Tokens, Visible Bills: The Urgent Need to Audit Hidden Operations in Opaque LLM Services. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2505.18471

Sun, N. X., Miao, Y., Jiang, H., Ding, M. D., & Zhang, J. (2024). From Principles to Practice: A Deep Dive into AI Ethics and Regulations. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2412.04683

Uysal, M. S. (2025). A multi-agent framework for verifiable AIGC licensing in digital ecosystems. *Gazi Iktisat ve Işletme Dergisi*, *11*(3). https://doi.org/10.30855/gjeb.2025.11.3.009

Wang, D., Liang, W., Chen, C., Xu, J., & Fu, Y. (2025). Governable AI: Provable Safety Under Extreme Threat Models. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2508.20411

Wenzel, J., Alam, S. U., Schmidt, A., Zhang, H., & Hermanns, H. (2025). A Workflow for Full Traceability of AI Decisions. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2511.11275

Wright, C. (2025). On Immutable Memory Systems for Artificial Agents: A Blockchain-Indexed Automata-Theoretic Framework Using ECDH-Keyed Merkle Chains. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2506.13246

Yanglet, X.-Y. L., Cao, Y., & Deng, L. (2025). Multimodal Financial Foundation Models (MFFMs): Progress, Prospects, and Challenges. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2506.01973

Zhang, H., Zhao, Y., Angione, C., Yang, H., Buban, J., Farhan, A., Johnston, F., & Colangelo, P. (2024). Towards Secure and Private AI: A Framework for Decentralized Inference. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2407.19401