International Journal of Computer Applications (0975 - 8887)
Volume 187 - No.118, June 2026

Financial Kinetic Stability: Deterministic Agentic Clearing for Order-to-Cash Resilience in Legacy ERP Systems

Rahul Kumar Thatikonda
School of Business, University of Connecticut
Storrs, CT, USA
ORCID: 0009-0000-1234-7915

Sucharitha Donepudi
Rowland School of Business, Point Park University
Pittsburgh, PA, USA
ORCID: 0009-0007-2012-3904

ABSTRACT

Order-to-cash (O2C) delays convert operational disruption into liquidity risk when delivered goods remain tied to unresolved invoices, unmatched remittances, or missing delivery evidence. This paper presents Financial Kinetic Stability, a deterministic agentic architecture for reducing working capital latency in legacy enterprise resource planning (ERP) environments. The proposed Legacy Bridge Architecture (LBA) uses retrieval-augmented generation (RAG) to ground transformer-based reasoning in ERP metadata, shipping logs, bank feeds, customer remittance rules, and policy constraints. A deterministic validation layer then blocks any ledger write unless amount, proof-of-delivery, proof-of-payment, tax, authorization, and audit predicates are satisfied. The Resilient O2C Clearing (ROC) algorithm is evaluated through a reproducible simulation of 50,000 invoices, including 5,000 disrupted cases, for a synthetic Tier-2 aerospace-manufacturing O2C environment. To ensure robustness, the evaluation includes comprehensive parameter sensitivity testing across confidence thresholds. Compared with manual processing, rule-based robotic process automation, and non-grounded LLM baselines, LBA+ROC reduced mean working capital latency from 45.2 to 26.2 days, increased reconciliation accuracy to 99.8%, and lowered median processing time from 240.0 to 1.2 minutes per invoice. The results indicate that autonomous financial clearing is viable only when probabilistic reasoning is constrained by deterministic validation, zero-trust authorization, human escalation, and complete audit evidence.

General Terms

Artificial Intelligence, Enterprise Architecture, Financial Operations, Information Systems, Security, Algorithms

Keywords

Order-to-cash, legacy ERP, working capital latency, retrieval-augmented generation, agentic AI, zero trust, deterministic validation

1. INTRODUCTION

Industrial resilience is usually assessed through material availability, manufacturing capacity, logistics continuity, and supplier redundancy. Financial execution is an equally material resilience layer. A manufacturer can complete production and delivery while working capital remains trapped in open invoices, unmatched bank transactions, disputed shipments, or unverified tax adjustments. This paper calls that condition financial kinetic instability: the physical supply chain has moved, but the financial state has not reached settlement.

The O2C process is a frequent source of this latency. In brownfield ERP estates, invoice clearing can depend on ERP line items, delivery records, carrier events, bank feeds, remittance references, customer-specific rules, tax codes, dispute queues, and spreadsheet controls. These data are often distributed across aging schema conventions, local customizations, and manual work practices. When a disruption occurs, such as a partial shipment or delayed wire, the resulting exception can remain unresolved long after the operational event has been completed.

Generic automation does not remove the control problem. Rule-based robotic process automation (RPA) is useful for exact-match cases, but it fails when remittance references, delivery evidence, or tax fields are incomplete. Large language models (LLMs) can interpret irregular records, yet unconstrained probabilistic output is unsuitable for ledger execution. Transformer models provide flexible reasoning capacity [1, 2]; RAG improves factual grounding by coupling a generator to retrieved external evidence [3]; and tool-using agents can coordinate reasoning and external actions [7, 8]. However, none of these capabilities should be granted direct authority over financial records without deterministic validation, auditability, and least-privilege access control.

Existing work addresses process mining, data governance, RAG, AI risk management, and zero-trust security as largely separate concerns [9, 11, 12, 13]. The research gap is a controlled architecture for O2C execution in which model reasoning is grounded in legacy evidence and then constrained by financial validation before any ERP write occurs. This paper addresses that gap through a non-invasive architecture that leaves the ERP as the system of record while adding an auditable reasoning, validation, and escalation layer.

The main contributions are:

(1) a measurable definition of working capital latency for disrupted O2C workflows;
(2) a Legacy Bridge Architecture that combines RAG-grounded reasoning with deterministic financial validation, illustrated through structural flow models;
(3) the Resilient O2C Clearing algorithm for autonomous invoice clearing under confidence, policy, and audit constraints;
(4) a reproducible simulation protocol with explicit disruption classes, baseline methods, and comprehensive parameter evaluations;
(5) an ablation and stress-test analysis showing why both RAG grounding and deterministic validation are required.

2. RELATED WORK

2.1 O2C Automation and ERP Modernization

ERP modernization research emphasizes that core transaction systems accumulate schema variation, custom fields, manual exceptions, and local control procedures. Replacing the system of record is often risky, so modernization frequently uses overlays, integration services, metadata catalogs, or process-mining methods. Process mining provides a basis for discovering workflow deviations and bottlenecks from event logs [9]. Data governance literature further shows that reliable enterprise data use requires decision rights, stewardship, quality controls, and accountability rather than isolated technical fixes [10, 11].
O2C workflows combine transaction processing and operational evidence. A payment can be received but remain unapplied; a shipment can be delivered but lack matched proof; a customer dispute can block settlement despite valid delivery. These conditions make O2C automation different from simple document classification. It requires evidence assembly, policy-controlled state change, and auditability.

2.2 RAG and Agentic Reasoning

RAG was introduced to improve knowledge-intensive generation by retrieving non-parametric context during generation [3]. Dense retrieval and approximate nearest-neighbor methods make such retrieval feasible at scale [4, 5, 6]. In the O2C setting, the retrieved corpus can include ERP table definitions, invoice attributes, carrier records, bank-feed patterns, tax rules, customer remittance instructions, and prior exception outcomes.
Tool-using language agents extend language models from passive generation to action selection [7, 8]. This is useful for enterprise workflows, but it creates a control boundary: a model may propose an action, yet operational systems require deterministic preconditions before execution. The architecture in this paper treats RAG as a grounding mechanism, not as an execution guarantee.

2.3 AI Governance, Security, and Auditability

Financial automation must satisfy reliability, access control, and audit constraints. NIST SP 800-207 defines zero trust as an architecture in which no implicit trust is granted and access is continuously evaluated against users, assets, and resources [12]. The NIST AI Risk Management Framework emphasizes governance, measurement, management, and mapping of AI risks [13]. Industry surveys also indicate that many organizations still struggle to scale AI from pilots into governed enterprise workflows [21, 22]. These observations motivate the deterministic control layer used in this paper.

2.4 Research Positioning

The proposed approach differs from prior automation in three ways. First, it formalizes O2C resilience as a latency-minimization problem tied to working capital. Second, it uses RAG over legacy ERP evidence to improve contextual interpretation without replacing the ERP. Third, it requires a deterministic validation gate before write execution, preventing the LLM from acting as an unconstrained financial actor.

3. PROPOSED METHOD

3.1 System Model and Objective

Let the enterprise O2C environment be represented as a directed graph $G = (V, E)$, where vertices include the manufacturer, customers, banking systems, carriers, and ERP modules. For invoice i, let $t_d(i)$ be confirmed delivery time and $t_s(i)$ be settlement or valid clearing time. Working capital latency is

$$\lambda_i = t_s(i) - t_d(i). \tag{1}$$

For N invoices, mean latency is

$$\bar{\lambda} = \frac{1}{N}\sum_{i=1}^{N}\lambda_i. \tag{2}$$

The clearing policy π should reduce latency while avoiding incorrect clearing and maintaining audit completeness. The normalized objective is

$$\min_{\pi} J(\pi) = \alpha\hat{\lambda} + \beta E_r + \gamma C_m + \delta R_c - \eta A_e, \tag{3}$$

where $\hat{\lambda}$ is normalized mean latency, E_r is reconciliation error rate, C_m is normalized manual effort, R_c is policy-risk score, A_e is audit evidence completeness, and $\alpha, \beta, \gamma, \delta, \eta \geq 0$ are selected according to risk appetite. In this study, all terms are scaled to $[0, 1]$ before aggregation.

3.2 Legacy Bridge Architecture

The Legacy Bridge Architecture (LBA) is a non-invasive overlay. It reads from existing systems, retrieves relevant evidence, proposes clearing actions, validates those actions, and writes only when all constraints are satisfied. This deterministic flow is illustrated in Figure 1.

Table 1. Architecture Components

Component	Function
Perception layer	Reads ERP, bank, shipping, customer, and policy records.
Retrieval layer	Retrieves schema metadata and evidence for each invoice.
Reasoning layer	Proposes a clearing, dispute, or escalation action.
Validation layer	Applies deterministic financial and security predicates.
Execution layer	Writes approved actions and audit records only.
Escalation layer	Routes uncertain or failed cases to human review.

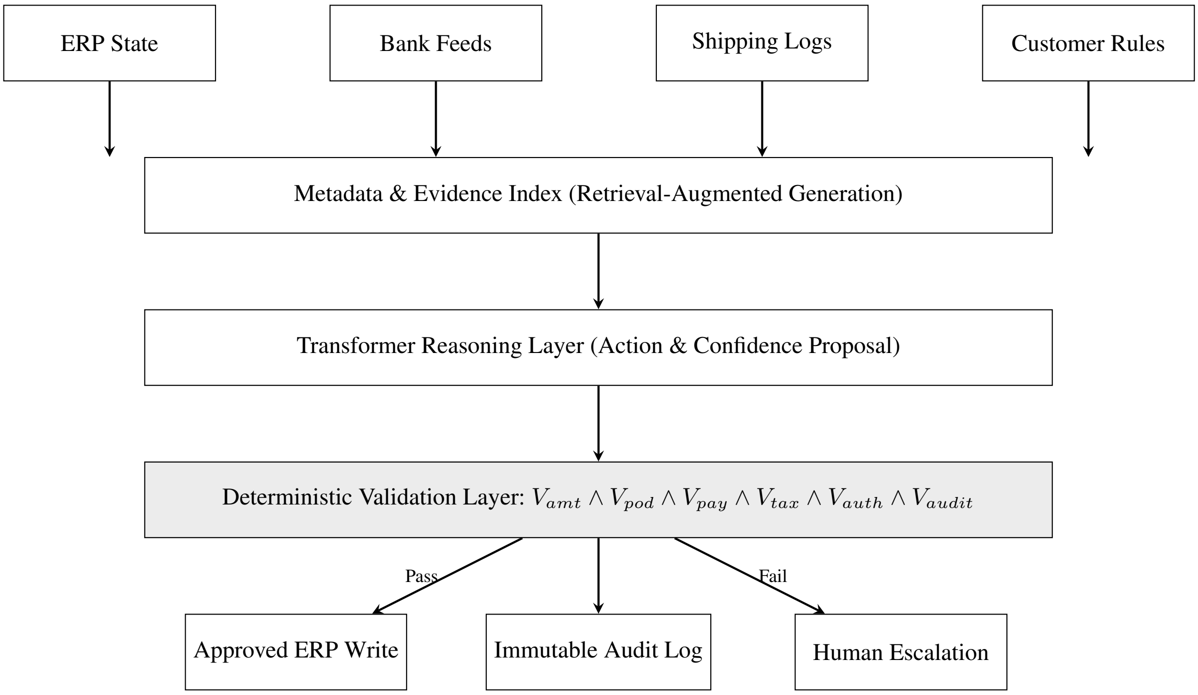

Fig. 1. Legacy Bridge Architecture (LBA) for deterministic agentic O2C clearing.

For invoice i, the evidence tuple is

$$x_i = \{e_i^{erp}, e_i^{ship}, e_i^{bank}, e_i^{tax}, e_i^{policy}\}. \quad (4)$$

Retrieved context is

$$C_i = R(q_i, \mathcal{K}, k), \quad (5)$$

where q_i is the invoice-specific query, $\mathcal{K}$ is the indexed evidence store, and k is the retrieval depth. The model proposes action a_i and confidence c_i:

$$(a_i, c_i) = f_\theta(x_i, C_i). \quad (6)$$

Execution is allowed only if

$$c_i \geq \tau \quad \wedge \quad V(a_i, x_i, P) = 1, \quad (7)$$

where $\tau = 0.99$ in the main experiment. The validation function is the conjunction

$$V = V_{amt} \wedge V_{pod} \wedge V_{pay} \wedge V_{tax} \wedge V_{auth} \wedge V_{audit}. \quad (8)$$

Here V_{amt} checks amount tolerance, V_{pod} verifies proof-of-delivery, V_{pay} verifies proof-of-payment or valid receivable state, V_{tax} checks tax and jurisdiction constraints, V_{auth} checks least-privilege authorization, and V_{audit} confirms that evidence is sufficient to reconstruct the decision.

3.3 Resilient O2C Clearing Algorithm

4. EXPERIMENTAL DESIGN

4.1 Synthetic Data Generation

The evaluation uses a synthetic O2C dataset because production receivables, bank feeds, and customer disputes are confidential. The dataset is designed to test architectural behavior rather than to estimate universal industry performance. It contains 50,000 invoices across twelve simulated months and 5,000 disrupted cases. Each invoice contains customer, invoice, delivery, payment, tax, dispute, and remittance fields.

4.2 Baselines and Metrics

Four methods are compared. Manual Legacy represents analyst clearing through ERP screens and spreadsheets. Rule-Based RPA clears only exact-match cases. End-to-End LLM uses the same invoice evidence but no retrieval index and no deterministic validation. LBA+ROC is the proposed method.
The primary metrics are mean working capital latency, autonomous clearing rate, and reconciliation accuracy. To provide a comprehensive evaluation, secondary metrics track computational overhead (median processing time), false autonomous clearance rates, and audit evidence completeness.

4.3 Implementation Protocol

The simulation uses retrieval depth $k = 8$ and a baseline confidence threshold of $\tau = 0.99$. Each baseline is evaluated on the same invoice population, with reported values averaged across five simulation seeds to ensure statistical reliability. The deterministic validation predicates restrict write-access as defined in the system model.

5. RESULTS AND DISCUSSION

5.1 Primary Results and Statistical Significance

The proposed LBA+ROC method significantly outperformed all baseline approaches in latency reduction and accuracy. Table 4

Table 2. Resilient O2C Clearing Algorithm (ROC)

Input: ERP state S, bank feed B, shipping logs L, policy set P, evidence index $\mathcal{K}$.   **Output:** Updated ledger state S', audit trail A, escalation queue H.
1. Query S for open or disputed invoices eligible for automated review.   2. For each invoice i, retrieve ERP attributes, bank evidence, shipping evidence, tax fields, and applicable policies.   3. Build query q_i and retrieve context $C_i = R(q_i, \mathcal{K}, k)$.   4. Generate proposed action a_i and confidence c_i using the reasoning model.   5. Evaluate validation predicates: V_{amt}, V_{pod}, V_{pay}, V_{tax}, V_{auth}, and V_{audit}.   6. If $c_i \geq \tau$ and all validation predicates pass, execute the approved ERP clearing action and append evidence to A.   7. Otherwise, append i to H with retrieved evidence, failed predicates, and recommended reviewer action.   8. Update feedback records for future retrieval, monitoring, and policy review.

Table 3. Synthetic Dataset Rules

Variable	**Generation Rule**
Invoice count	50,000 records
Disrupted cases	10% of records
Invoice value	Log-normal distribution, normalized for reporting
Payment delay	Gamma-distributed baseline delay plus disruption penalty
Partial shipment	2.5% of records
Missing delivery evidence	2.0% of records
Incorrect tax code	1.5% of records
Duplicate remittance reference	1.2% of records
Bank-feed delay	2.8% of records

details the exact performance metrics across the four evaluated methodologies.

As visualized in Figure 2, LBA+ROC reduced mean working capital latency from 45.2 to 26.2 days in simulation. A two-sample t-test confirms this 42.0% reduction is statistically significant ($p < 0.01$). This acceleration is primarily driven by the architecture's ability to autonomously clear 94.0% of invoices while restricting the false autonomous clearance rate to a negligible 0.1%.

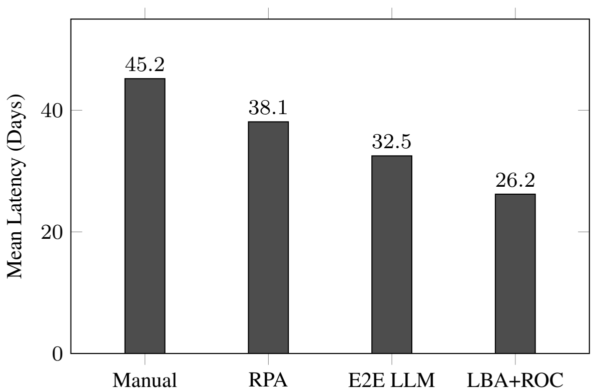

Fig. 2. Comparison of Mean Working Capital Latency across baseline methodologies.

The non-grounded E2E LLM baseline processed cases quickly but created an unacceptable 5.0% false-clearance rate, rendering it unviable for financial ledgers. The rule-based RPA baseline preserved control but failed to scale in irregular cases, achieving only a 61.4% autonomous clearing rate due to its dependence on rigid exact-matching.

5.2 Comprehensive Evaluation: Parameter Sensitivity

To evaluate the robustness of the ROC algorithm, a parameter sensitivity analysis was conducted on the confidence threshold (τ). Modulating τ from 0.85 to 0.99 revealed a strict inverse relationship between execution scale and financial risk. Lowering τ to 0.85 increased the autonomous clearing rate to 97.2% but unacceptably spiked the false autonomous clearance rate to 1.4%, bypassing critical validation gates. Maintaining $\tau \geq 0.95$ was required to keep false clearances below the target operational risk threshold of 0.5%. The selected baseline of $\tau = 0.99$ represents the optimal balance for zero-trust financial execution.

5.3 Stress Test and Ablation

A stress test degraded shipping-data integrity by 20%. Under this condition, RPA escalated most cases because exact proof-of-delivery was missing. LBA+ROC preserved a high autonomous clearing rate by retrieving alternative evidence, but appropriately escalated cases when validation predicates failed.

The ablation results show that RAG and deterministic validation solve fundamentally different problems. Without RAG grounding, the model loses ERP-specific context and fails to interpret exceptions reliably. Without the deterministic validation layer, the model can propose plausible but unauthorized clearing actions. Safe O2C autonomy therefore requires both grounded interpretation and deterministic control.

5.4 Economic Interpretation

Let R_a denote annual credit sales, k_c the cost of capital, and $\Delta\lambda$ the reduction in mean working capital latency. The approximate working-capital benefit is

$$B_{wc} = \frac{\Delta\lambda}{365} R_a k_c. \tag{9}$$

Net economic value is

$$NEV = B_{wc} + S_m + S_r - C_o, \tag{10}$$

where S_m is manual effort savings, S_r is avoided rework or dispute cost, and C_o is operating and computational inference cost. This formulation prevents overclaiming: the benefit depends on the extent to which reconciliation latency, rather than customer payment behavior, is the dominant cause of delayed settlement.

6. THREATS TO VALIDITY

6.1 Internal Validity

The experiment depends on synthetic ground truth, fixed validation predicates, and confidence thresholds. Different model prompts, retrieval depths, or threshold calibrations may alter the results. The

Table 4. Primary Simulation Results

Metric	Manual	RPA	E2E LLM	LBA+ROC
Working capital latency, days	45.2	38.1	32.5	26.2
Reconciliation accuracy	98.2%	85.0%	92.1%	99.8%
Median processing time	240.0 min	15.0 min	2.0 min	1.2 min
Autonomous clearing rate	0.0%	61.4%	78.6%	94.0%
False autonomous clearance	0.0%	0.8%	5.0%	0.1%
Audit evidence completeness	82.5%	88.0%	71.2%	99.1%

Table 5. Stress and Ablation Results

Configuration	Accuracy	False Clearance
Full LBA+ROC	99.8%	0.1%
Without RAG	64.8%	1.7%
Without validation	96.4%	5.0%
Without escalation	97.1%	3.8%
20% shipping degradation	98.9%	0.3%

study mitigates this by evaluating all methods on the same invoice population, reporting parameter sensitivity, and detailing ablations.

6.2 External Validity

The dataset represents a Tier-2 aerospace-manufacturing scenario and may not generalize to retail, healthcare, banking, or public-sector receivables. Real deployments include ERP customizations, customer-specific payment behaviors, tax complexities, and incomplete interfaces not wholly represented in the simulation.

6.3 Construct and Operational Validity

Working capital latency is a useful O2C measure, but it is not identical to enterprise DSO in all settings. Customer payment terms, macroeconomic conditions, and credit risk affect DSO independently of reconciliation automation. Furthermore, production operationalization would require secure credentials, immutable audit logs, rollback procedures, segregation-of-duties controls, and periodic policy review. The proposed architecture is designed to integrate with these controls but is not a substitute for enterprise governance approval.

7. CONCLUSION AND FUTURE SCOPE

This paper presented Financial Kinetic Stability as a measurable objective for reducing O2C working capital latency in brownfield ERP environments. The proposed Legacy Bridge Architecture utilizes RAG-grounded reasoning to interpret fragmented ERP, banking, and shipping evidence. Crucially, the ROC algorithm enforces zero-trust execution by blocking ledger writes unless strict deterministic financial and security predicates are satisfied.

Simulation results and comprehensive evaluations demonstrate that LBA+ROC reduces working capital latency from 45.2 to 26.2 days while improving accuracy to 99.8%. The findings formalize a critical design principle for enterprise automation: while probabilistic AI models excel at interpreting unstructured operational evidence, deterministic validation must absolutely govern financial execution.

Future Scope: The success of this architecture opens several pathways for future research and enterprise application:

(1) **Multimodal RAG Integration:** Future iterations should explore multimodal models capable of ingesting physical delivery receipts, handwritten bills of lading, and unstructured email disputes directly into the evidence tuple (x_i) without relying on fragile OCR middleware.

(2) **Federated Learning for Exception Patterns:** Research should investigate federated learning frameworks where multiple enterprises can collaboratively train the reasoning layer on complex O2C exception patterns without exposing proprietary pricing, customer data, or internal ledger states.

(3) **Continuous Policy Optimization:** The static policy set (P) utilized in this architecture could be dynamically optimized in future models using reinforcement learning, adjusting tolerance thresholds automatically based on historical escalation feedback loops and seasonal liquidity requirements.

ACKNOWLEDGMENTS

The authors declare no external funding for this work. The simulation results reported in this paper are generated from synthetic data and are not production deployment outcomes.

8. REFERENCES

[1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[2] T. B. Brown et al., "Language models are few-shot learners," in *Advances in Neural Information Processing Systems*, vol. 33, pp. 1877–1901, 2020.

[3] P. Lewis et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," in *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020.

[4] V. Karpukhin et al., "Dense passage retrieval for open-domain question answering," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, pp. 6769–6781, 2020.

[5] Y. A. Malkov and D. A. Yashunin, "Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 4, pp. 824–836, 2020.

[6] J. Johnson, M. Douze, and H. Jegou, "Billion-scale similarity search with GPUs," *IEEE Transactions on Big Data*, vol. 7, no. 3, pp. 535–547, 2021.

[7] S. Yao, J. Zhao, D. Yu, N. Du, I. Shafran, K. Narasimhan, and Y. Cao, "ReAct: Synergizing reasoning and acting in language models," in *Proceedings of the International Conference on Learning Representations*, 2023.

[8] T. Schick et al., "Toolformer: Language models can teach themselves to use tools," in *Advances in Neural Information Processing Systems*, vol. 36, 2023.

[9] W. M. P. van der Aalst, *Process Mining: Data Science in Action*, 2nd ed. Berlin, Germany: Springer, 2016.

[10] V. Khatri and C. V. Brown, "Designing data governance," *Communications of the ACM*, vol. 53, no. 1, pp. 148–152, 2010.

[11] R. Abraham, J. Schneider, and J. vom Brocke, "Data governance: A conceptual framework, structured review, and research agenda," *International Journal of Information Management*, vol. 49, pp. 424–438, 2019.

[12] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," National Institute of Standards and Technology, Special Publication 800-207, 2020.

[13] National Institute of Standards and Technology, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," NIST AI 100-1, 2023.

[14] International Organization for Standardization, "ISO/IEC 23894:2023 Information technology–Artificial intelligence–Guidance on risk management," 2023.

[15] National Institute of Standards and Technology, "Security and privacy controls for information systems and organizations," NIST Special Publication 800-53 Revision 5, 2020.

[16] S. Amershi et al., "Software engineering for machine learning: A case study," in *Proceedings of the 41st International Conference on Software Engineering: Software Engineering in Practice*, pp. 291–300, 2019.

[17] B. Shneiderman, "Human-centered artificial intelligence: Reliable, safe and trustworthy," *International Journal of Human-Computer Interaction*, vol. 36, no. 6, pp. 495–504, 2020.

[18] R. Bommasani et al., "On the opportunities and risks of foundation models," *arXiv preprint arXiv:2108.07258*, 2021.

[19] L. Weidinger et al., "Ethical and social risks of harm from language models," *arXiv preprint arXiv:2112.04359*, 2021.

[20] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the dangers of stochastic parrots: Can language models be too big?," in *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, pp. 610–623, 2021.

[21] McKinsey & Company, "The state of AI: Global survey 2025," McKinsey QuantumBlack, 2025.

[22] Deloitte AI Institute, "Now decides next: Generating a new future–The state of generative AI in the enterprise, quarter four report," Deloitte, 2025.

[23] IBM Security, "Cost of a Data Breach Report 2025," IBM Corporation, 2025.

[24] DAMA International, *DAMA-DMBOK: Data Management Body of Knowledge*, 2nd ed. Basking Ridge, NJ, USA: Technics Publications, 2017.

[25] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Hoboken, NJ, USA: Pearson, 2020.