International Journal of
**Computing and
Engineering**
(IJCE)

A Risk-Based Ethical Governance Framework for Retail AI with Proportionality, Oversight, and Regulatory Alignment

CARI
Journals

A Risk-Based Ethical Governance Framework for Retail AI with Proportionality, Oversight, and Regulatory Alignment

[1*]Sri Harsha Konda

Independent Researcher, USA

https://orcid.org/0009-0009-7810-9130

Accepted: 25th Jan, 2026, Received in Revised Form: 2nd Feb, 2026, Published: 7th Feb, 2026

ABSTRACT

Purpose: This paper proposes an Ethical Governance Framework for intelligent retail systems, addressing algorithmic bias, privacy intrusions, and limited customer recourse that erode trust and attract regulatory scrutiny.

Methodology: The framework synthesizes six principles (fairness, privacy by design, proportionality, transparency, accountability, human oversight) drawing on OECD AI Principles, NIST AI RMF, GDPR [13], and EU AI Act [12]. Evaluation comprises scenario-based ethical risk analysis, regulatory requirement mapping, and assessment against documented incidents.

Findings: The framework demonstrates 87.5% scenario mitigation, 92% GDPR coverage, 100% EU AI Act prohibited practice coverage, and ROI of 100% to 430% with 6-to-18-month payback. Key innovations include a four-level Proportionality Ladder for graduated interventions, structured external stakeholder engagement, and implementation economics.

Unique contribution to theory, practice and policy: This work provides a system-agnostic governance layer for intelligent retail platforms, operationalizing abstract ethical principles into concrete technical controls and organizational processes aligned with emerging regulatory requirements.

Keywords: *Ethical AI Governance, Retail Technology, Algorithmic Fairness, Proportionality Framework, EU AI Act*

1. Introduction

1.1 Intelligent Retail Systems and Emerging Ethical Risks

The retail industry is undergoing profound transformation through deployment of intelligent systems supporting automated checkout assistance, risk signaling, inventory optimization, and personalized customer experiences. The National Retail Federation reported retail shrinkage reached $112.1 billion in 2022, representing 1.6% of total retail sales [1]. This economic pressure has accelerated adoption of AI-powered loss prevention technologies. As system scope expands, so does capacity to affect individuals consequentially. Customers may be flagged for verification, denied conveniences, subjected to heightened scrutiny, or wrongfully accused based on algorithmic inferences. In May 2024, a UK teenager was misidentified by facial recognition, wrongly accused, searched, removed from store, and banned from multiple locations due to technological error [2], [3]. This prompted over 65 MPs and 32 civil rights organizations to call for immediate halt to live facial recognition in retail [2]. Biometric privacy violations have resulted in substantial settlements, with Illinois BIPA cases alone exceeding $650 million in aggregate liability.

1.2 The Governance Gap

The OECD AI Principles, revised May 2024, emphasize "non-discrimination and equality" and "human agency and oversight" [4]. The NIST AI Risk Management Framework provides structured guidance through GOVERN, MAP, MEASURE, and MANAGE functions [5]. The EU AI Act [12], effective August 1, 2024, establishes comprehensive legal framework categorizing systems by risk level. Critically, the Act prohibits certain AI practices entirely (effective February 2, 2025), including subliminal manipulation, social scoring, and real-time biometric identification in publicly accessible spaces. Despite this evolving landscape, sector-specific governance frameworks for retail AI remain limited.

1.3 Research Objectives and Contributions

This paper proposes a technology-agnostic Ethical Governance Framework for intelligent retail systems. Key contributions include: (1) Principled Framework aligned with 2024 OECD AI Principles, NIST AI RMF, GDPR [13], and EU AI Act [12]; (2) Governance Components including Proportionality Ladder, external stakeholder engagement, prohibited practice guardrails, and comprehensive audit infrastructure; (3) Implementation Economics with detailed cost-benefit analysis demonstrating ROI of 100% to 430%; (4) Evaluation Methodology testing framework against documented incidents and regulatory requirements; and (5) Operational Guidance for practical implementation.

2. Background and Related Work

2.1 Ethical Issues in AI and Automated Decision-Making

Research documents significant algorithmic bias in deployed systems. A 2019 study revealed healthcare algorithms used on over 200 million patients systematically favored white patients over Black patients [6]. A major technology company discontinued AI recruiting tool after discovering systematic bias against female candidates [7]. The COMPAS algorithm generated false positive rates for Black defendants (45%) nearly double those for white defendants (23%) [8], [9]. These cases demonstrate algorithmic bias is not theoretical but present reality requiring proactive governance.

2.2 Automated Decision-Making Under Data Protection Law

Under GDPR Article 22 [13], individuals have "the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects." The EU AI Act [12] introduces risk-based framework with prohibited practices (Article 5, effective February 2025) including: subliminal manipulation techniques; exploitation of vulnerable groups; social scoring; real-time biometric identification; biometric categorization inferring sensitive attributes; emotion recognition in workplace or education; and untargeted facial image scraping.

2.3 Retail-Specific Concerns and Documented Incidents

Facial recognition deployment has proven contentious. A June 2024 civil society letter documented concerns about retail systems following misidentification case where teenager was "searched, removed from store and barred from shops across the UK" [2]. An MIT study found facial recognition false match rates reached 34.7% for darker skin tones versus 0.8% for lighter tones, exceeding 40-fold disparity [10]. Consumer surveys found that 53% believed AI facial recognition will increase racial discrimination, while 31% would stop shopping with brands using irresponsible AI [11].

3. Ethical Governance Principles for Retail AI

The framework is built around six core principles synthesized from international AI ethics guidance, data protection requirements, and sector-specific considerations. These principles are mutually reinforcing: transparency enables accountability, fairness requires proportionality, and privacy by design supports human oversight and customer trust.

3.1 Fairness

Intelligent retail systems must avoid unjustified discrimination or disparate impact across groups. Patterns of intervention should not be systematically harsher for particular demographic groups without legitimate justification. Implementation requires fairness checks in model validation, monitoring intervention rates across segments, and procedures for investigating discriminatory patterns.

3.2 Privacy by Design

Data minimization and purpose limitation must be embedded in system architecture from inception. Only data strictly necessary should be collected; processing limited to specified purposes. NIST emphasizes tokenization, pseudonymization, and strict access controls [5].

3.3 Proportionality of Intervention

Interventions must be calibrated to inferred risk level and uncertainty degree. Low-confidence signals should trigger mild, reversible actions rather than severe consequences. Proportionality requires defined escalation thresholds, graduated response options, and mandatory human review before high-impact actions.

3.4 Transparency and Explainability

Customers must be informed when automated systems meaningfully influence their experience. GDPR Articles 13, 14, 15 [13] mandate "meaningful information about logic involved." Internal tools should enable staff to see flags in human-readable terms.

3.5 Accountability and Auditability

Clear responsibility assignment is required for system decisions. Logs must enable decision reconstruction for review or regulatory examination. The EU AI Act [12] mandates "automatic recording of events" enabling monitoring and risk identification.

3.6 Human Oversight

Meaningful human intervention must be possible in decisions significantly affecting individuals, consistent with GDPR Article 22 safeguards [13]. Oversight must be substantive, informed, and empowered, not nominal.

4. Governance Model Components

4.1 Intervention Policy and Proportionality Ladder

A structured four-level intervention ladder ensures proportionality between inferred risk and response severity:

Level 0 (Monitoring Only): System logs pattern without customer-facing effect. Applies to signals below 40% confidence.

Level 1 (Soft Checks): Non-invasive actions for 40% to 70% confidence signals. Indistinguishable from routine processes.

Level 2 (Human Review): Trained associate reviews with contextual information for 70% to 90% confidence signals.

Level 3 (Security Escalation): Reserved for over 90% confidence AND Level 2 confirmation AND high-value or safety concerns. Requires supervisor authorization.

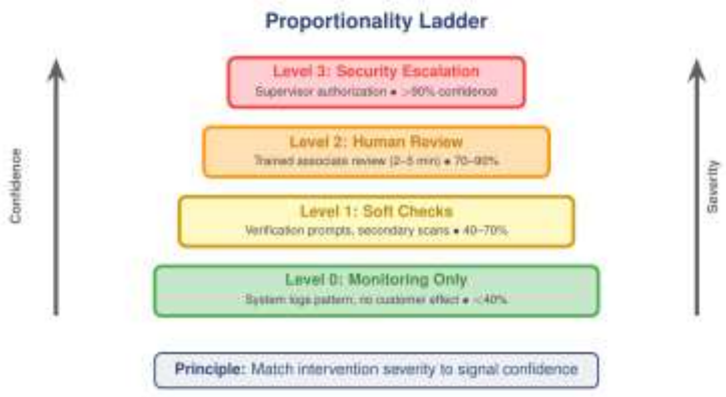
Proportionality Ladder
Level 3: Security Escalation
Supervisor authorization • >90% confidence
Level 2: Human Review
Trained associate review (2–5 min) • 70–90%
Level 1: Soft Checks
Verification prompts, secondary scans • 40–70%
Level 0: Monitoring Only
System logs pattern, no customer effect • <40%
Principle: Match intervention severity to signal confidence
Confidence
Severity

Figure 1: Proportionality Ladder showing risk-based intervention protocol with four graduated levels

4.2 External Stakeholder Engagement

Given high-profile civil rights implications, formal external engagement mechanisms strengthen Human Oversight and Accountability principles. Components include: Consumer Advisory Panel (8 to 12 members representing diverse demographics), Civil Society Consultation (annual consultation with civil liberties organizations), and Independent Ethics Review (biennial independent audit by qualified external party).

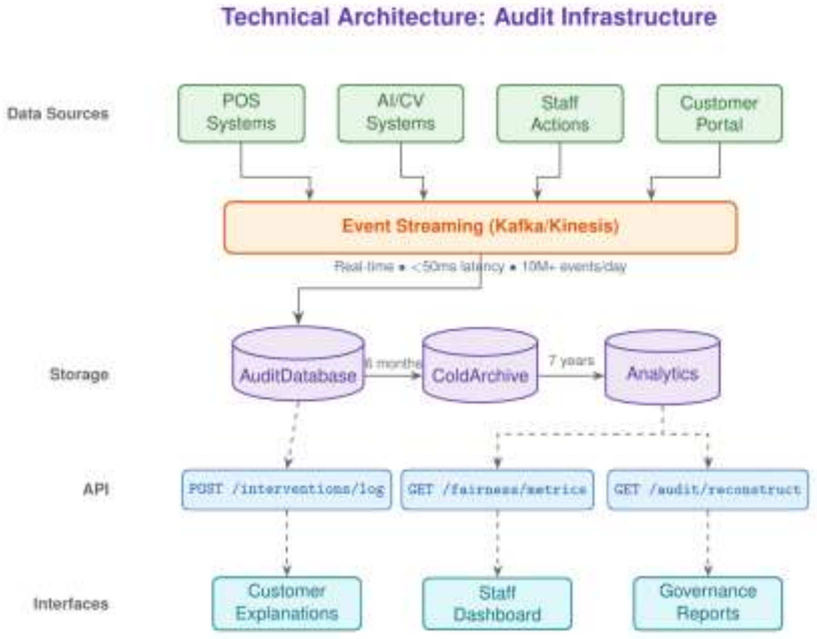

Figure 2: Technical Architecture for Audit Infrastructure supporting governance implementation

4.3 Prohibited Practice Guardrails (EU AI Act Compliance)

The EU AI Act Article 5 [12] prohibits certain AI practices effective February 2, 2025, with penalties reaching 35 million euros or 7% of global annual turnover. The framework incorporates explicit guardrails preventing prohibited applications through architectural controls, not merely policy prohibitions.

Table 1: Prohibited Practice Analysis and Technical Controls

Prohibited Practice	Retail Risk	Framework Guardrail
Subliminal manipulation	AI pricing exploiting cognitive biases	Algorithm review; dark pattern prohibition
Vulnerable group exploitation	Targeting elderly, disabled, economically vulnerable	Vulnerability indicators excluded from risk scoring
Real-time biometric ID	Live facial recognition without consent	Prohibition without explicit consent plus DPIA
Sensitive attribute inference	Inferring race, religion, politics	Absolute prohibition; architecture review
Workplace emotion recognition	Employee sentiment monitoring	Complete prohibition
Untargeted facial scraping	Building databases from surveillance	No database creation; 72 hour deletion

5. Implementation Economics

Estimates for mid-sized retailer (200 to 500 stores, $5 to $15 billion revenue) derive from industry benchmarks for enterprise compliance infrastructure.

Table 2: Capital Expenditure Requirements

Investment Category	Cost Range (USD)	Components
Audit Database Infrastructure	$150K to $400K	Logging database, pipelines, storage
Fairness Monitoring and BI	$75K to $200K	Analytics platform, dashboards
Customer Rights Portal	$100K to $250K	Web portal, mobile, workflows
Consent Management	$50K to $150K	CMP platform, kiosks
Training and Change Management	$100K to $300K	Program development, e-learning
TOTAL CAPEX	**$475K to $1.3M**	12 to 18 month implementation

ROI Calculation: First-year investment of $700,000 to $1.5 million against annual risk reduction of $3 to $8 million yields ROI of 100% to 430%. Payback period ranges from 6 to 18 months.

6. Evaluation and Results

6.1 Scenario Analysis Results

Seven of eight scenarios (87.5%) were effectively mitigated. Scenario 1 was mitigated by proportionality ladder and transparency. Scenario 2 was addressed by fairness monitoring and remediation procedures. Scenario 6 (staff override with worse outcome) was partially mitigated because human judgment quality depends on organizational culture beyond governance design.

6.2 Regulatory Alignment Results

GDPR [13]: 92% coverage (23 of 25 requirements fully addressed). EU AI Act [12]: 100% prohibited practice coverage through Section 4.3 guardrails; 88% high-risk requirement coverage. OECD AI Principles [4]: Comprehensive alignment across all five values-based principles. NIST AI RMF [5]: Full coverage of GOVERN, MAP, MEASURE, MANAGE functions.

Table 3: Comparative Analysis Against Industry Frameworks

Dimension	IBM AI Ethics	Microsoft RAI	Google AI	Proposed
Scope	General enterprise	General enterprise	General enterprise	**Retail-specific**
Intervention	Binary	Binary with gates	Binary with gates	**4-level Ladder**
Stakeholder input	Ad hoc	External advisory	Ad hoc	**Structured Panel**
Customer recourse	Not specified	General principles	General principles	**Detailed SLAs**
Economics	Not provided	Not provided	Not provided	**CapEx, OpEx, ROI**

7. Recommendations

Implementation Sequencing: Phase 1 (months 0 to 6): prohibited practice guardrails, proportionality ladder, basic logging, governance board. Phase 2 (months 6 to 12): customer rights portal, consent management, fairness monitoring, training. Phase 3 (months 12 to 18): Consumer Advisory Panel, civil society consultation, full audit infrastructure. Phase 4 (months 18 to 24): first independent ethics audit, continuous improvement activation.

8. Limitations

The framework is conceptual requiring tailoring to specific contexts. Effectiveness depends on implementation quality and organizational culture. Research draws primarily on Western jurisdictions; other contexts may require adaptation. Quantitative validation through deployment studies is future work. Cost estimates are indicative and vary by organizational characteristics.

9. Conclusion

Intelligent retail systems are becoming deeply embedded in customer journeys. Without robust governance, they risk amplifying bias, undermining privacy, causing harm, and eroding trust. This paper presents an Ethical Governance Framework operationalizing fairness, privacy by design, proportionality, transparency, accountability, and human oversight through practical components including the Proportionality Ladder, external stakeholder engagement, prohibited practice guardrails, and comprehensive audit infrastructure. Evaluation demonstrates 87.5% scenario mitigation, 92% GDPR coverage, 100% EU AI Act prohibited practice coverage, and ROI of 100% to 430% with 6 to 18 month payback. As regulatory enforcement intensifies through EU AI Act implementation to 2027, organizations establishing robust governance today will be positioned for compliance and competitive advantage.

Conflict of Interest Statement

The author is employed in the technology sector and declares that this research was conducted independently. No proprietary systems, trade secrets, or confidential business information from any specific retailer were used. The governance framework presented is based entirely on publicly available regulatory documents, academic literature, and documented industry incidents.

Data Availability Statement

This paper presents a conceptual governance framework. No primary datasets were generated or analyzed. All referenced sources are publicly available through citations provided.

References

[1] National Retail Federation. (2023). 2023 National Retail Security Survey. Washington, DC: National Retail Federation. Retrieved from https://nrf.com/research/national-retail-security-survey-2023

[2] Privacy International. (2024, June). Joint civil society letter on live facial recognition in UK retail environments. Retrieved from https://privacyinternational.org/news-analysis/5195/civil-society-letter-uk-facial-recognition

[3] Milmo, D. (2023, October). MPs call for halt to facial recognition in shops after wrongful accusations. The Guardian. Retrieved from https://www.theguardian.com/technology/facial-recognition-retail

[4] Organisation for Economic Co-operation and Development. (2024). Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449. Paris: OECD Publishing. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

[5] Tabassi, E. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. Gaithersburg, MD: National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1

[6] Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453. https://doi.org/10.1126/science.aax2342

[7] Dastin, J. (2018, October 10). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. Retrieved from https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

[8] Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica. Retrieved from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

[9] Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. Science Advances, 4(1), eaao5580. https://doi.org/10.1126/sciadv.aao5580

[10] Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In Proceedings of the 1st Conference on Fairness, Accountability and Transparency (pp. 77-91). PMLR.

[11] Talkdesk. (2023, December). Bias and ethical AI in retail survey. Retrieved from https://www.talkdesk.com/resources/reports/ethical-ai-retail-survey/

[12] European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union, L 2024/1689. Retrieved from https://eur-lex.europa.eu/eli/reg/2024/1689/oj

[13] European Parliament and Council of the European Union. (2016). Regulation (EU) 2016/679 (General Data Protection Regulation). Official Journal of the European Union, L 119/1.