Computer Fraud and Security
ISSN (online): 1873-7056

Agentic Commerce at Scale: A Reference Architecture for Enterprise Retail Systems

Prakash Kodali

Sri Venkateswara University, India

Abstract

Autonomous shopping and operations through specialized agent networks represent the next evolution in retail, yet most implementations remain constrained to experimental deployments due to fundamental gaps in safety, governance, and enterprise integration. This article presents a comprehensive cloud-native reference architecture that enables production-grade agentic commerce across the retail technology stack. The architecture encompasses buyer-side agents for discovery, negotiation, and checkout operating at the edge and in cloud environments, alongside merchant-side agents for catalog intelligence, dynamic promotions, and service automation, all governed by policy-enforced guardrails. A novel orchestration layer provides model routing, tool access control, memory management, and multi-stage safety filtering, while an integrated observability plane tracks reliability, cost, and risk metrics aligned to service-level objectives. The framework addresses critical enterprise constraints, including PII protection, brand safety, regulatory compliance, and seamless integration with existing product information management, order management, inventory, and content systems. Three production case studies demonstrate practical deployment patterns for promotion negotiation, returns automation, and catalog question-answering, establishing validated pathways from experimental prototypes to enterprise-scale autonomous commerce platforms.

Keywords: Agentic commerce, autonomous agents, enterprise e-commerce architecture, retail automation, safety guardrails

1. Introduction & Landscape

1.1 Motivation and the Deployment Gap

Autonomous shopping agents represent an evolving paradigm within retail technology, presenting considerable opportunities for transforming product discovery, transaction optimization, and purchase completion through minimal human intervention. Recent innovations in retail automation, notably intelligent cart systems and self-navigating shopping frameworks, demonstrate meaningful advancement toward frictionless consumer experiences [1][2]. Despite these technological gains, operational implementations of agentic commerce platforms remain predominantly constrained within experimental settings and validation prototypes. Substantial barriers exist between laboratory demonstrations and industrial-strength deployments, stemming largely from unresolved challenges in safety assurance, compliance verification, and unified integration with legacy enterprise systems, including Product Information Management platforms, Order Management Systems, Inventory Management Systems, Content Management Systems, and Retail Media Networks. Large-scale retail organizations face a critical decision point where underlying technologies enabling autonomous agents have reached technical viability through developments in large language models and augmented computational frameworks, while the corresponding operational infrastructure necessary for commercial-scale implementation remains underdeveloped. This architectural framework addresses the identified implementation gap through a systematic cloud-native design intended to bridge theoretical agent functionalities with practical organizational requirements.

1.2 Scope and Contributions

This architectural framework delivers multiple substantive advances for autonomous commerce platforms. The design establishes systematic categorization and technical specifications for agent responsibilities distributed across consumer-facing and retailer-facing operational contexts. Consumer-facing agents integrate discovery functions, extracting user preferences and matching them against product catalogs, negotiation capabilities refining price structures and bundle configurations, cart assembly logic ensuring compatibility and optimizing value propositions, and checkout coordination managing payment processing alongside fraud detection. Retailer-facing agents incorporate catalog interrogation systems grounded in product knowledge bases, promotional engines operating within defined policy boundaries, and automated

service handling for returns management, refund processing, and issue resolution. The architectural design introduces an orchestration layer coordinating agent interactions through policy enforcement mechanisms, ensuring autonomous operations maintain conformance with organizational standards, brand guidelines, and regulatory obligations. Measurement frameworks encompass reliability indicators, cost tracking dimensions, and risk monitoring parameters, enabling organizations to assess agent performance relative to service-level targets. The validation approach combines offline simulation using synthetic workloads with controlled online experimentation through structured testing protocols, establishing thorough verification procedures for agent behavior before commercial deployment.

1.3 Paper Organization

The following sections develop this architectural framework through a structured presentation. Section 2 provides foundational context for agentic patterns within commerce domains, outlines organizational constraints and threat classifications, defines functional specifications for consumer and retailer agents, and establishes performance indicators with termination criteria for automated fallback activation. Section 3 details the core architectural design, describing system topology and foundational principles, orchestration layer elements, consumer-facing and retailer-facing agent implementations, and layered safety with governance mechanisms. Section 4 investigates operational considerations, including observability frameworks, reliability strategies with degradation handling, performance optimization approaches, validation methodologies, and three deployment scenarios demonstrating practical implementation techniques. Section 5 evaluates current limitations, identifies open research directions, discusses standardization efforts and ecosystem development, and examines emerging capabilities anticipated in future iterations. Section 6 consolidates architectural contributions, provides actionable recommendations for production deployment, articulates future trajectories for autonomous commerce systems, and encourages collaborative progress across the field.

2. Background & Requirements Engineering

2.1 Agentic Patterns in Commerce

Commerce-oriented agentic systems function through recurring Plan-Act-Observe-Learn sequences that support adaptive reasoning within evolving retail environments. These operational sequences allow continuous behavior refinement driven by environmental responses and performance evaluation. Tool-augmented large language model configurations enhance agent functionalities beyond text generation by connecting external computational utilities, information repositories, and application programming interfaces [3]. This augmentation permits agents to perform sophisticated retail tasks, including stock verification, price computation, and payment orchestration, that transcend standalone language model capacities. Memory architectures and contextual persistence mechanisms sustain temporal consistency throughout prolonged shopping interactions, retaining customer inclinations, dialogue records, and transaction states essential for tailored engagement. Multi-agent collaboration frameworks define messaging conventions and workload allocation structures supporting cooperative challenge resolution across distributed agent ensembles, especially critical when consumer-facing and retailer-facing agents negotiate agreements or reconcile discrepancies within integrated commerce pipelines.

2.2 Enterprise Constraints and Threat Model

Brand protection and factual accuracy hazards represent foremost considerations for commercial implementation, given that generative agents might yield erroneous merchandise details, unsuitable suggestions, or communications diverging from brand tone and messaging guidelines [4]. Personal data protection and sensitive information governance requirements enforce rigorous restrictions on information gathering, archival, transmission, and lifecycle management to safeguard consumer particulars across agent engagements. Compliance mandates encompassing General Data Protection Regulation, California Consumer Privacy Act, and Payment Card Industry Data Security Standard necessitate particular technical and operational controls that agent frameworks must embed to sustain lawful functionality across geographical boundaries. Connectivity challenges with established retail technology stacks introduce considerable technical obstacles, given that historical Product Information Management platforms, Order Management Systems, Inventory Management Systems, Content Management Systems, and Retail Media Networks commonly utilize disparate information structures, interaction protocols, and verification approaches that agent coordination substrates must harmonize.

Enterprise System	Integration Purpose	Primary Challenges	Required Safeguards
Product Information Management (PIM)	Product catalog access and attribute retrieval	Schema heterogeneity, update propagation delays	Canonical data models, change data capture
Order Management System (OMS)	Transaction processing and order lifecycle tracking	Transaction consistency, rollback mechanisms	Distributed transaction protocols, idempotency guarantees
Inventory Management System (IMS)	Stock availability verification and reservation	Real-time accuracy, reservation conflicts	Optimistic locking, availability buffers
Content Management System (CMS)	Brand assets and messaging consistency	Content approval workflows, version control	Staged deployment, A/B testing frameworks
Retail Media Network (RMN)	Advertising integration and sponsored placements	Attribution tracking, conflict of interest	Transparency disclosures, ethical guardrails
Payment Gateways	Transaction authorization and settlement	PCI-DSS compliance, fraud detection	Tokenization, end-to-end encryption
Customer Data Platforms	Profile unification and preference storage	GDPR/CCPA compliance, consent management	Purpose limitation, data minimization, retention policies

Table 1: Enterprise Integration Requirements and Challenges [4][5]

2.3 Functional Requirements

Consumer-facing agents execute shopper-oriented capabilities encompassing discovery, negotiation, cart assembly, and checkout procedures. Discovery agents locate pertinent merchandise matching customer intentions through semantic exploration, preference extraction, and suggestion synthesis. Negotiation agents enhance acquisition terms by evaluating options, recognizing applicable incentives, and structuring product combinations. Cart assembly agents verify congruence among chosen items, confirm stock availability, and propose supplementary merchandise. Checkout agents coordinate payment execution, administer reductions, compute levies and delivery charges, and synchronize fraud screening operations. Retailer-facing agents bolster merchant workflows through catalog interrogation, adaptive promotions, and service mechanization features. Catalog interrogation agents deliver substantiated answers to merchandise queries utilizing knowledge repositories and organized product specifications. Adaptive promotion agents construct personalized incentives within governance limitations based on customer classifications, stock quantities, and commercial targets. Service mechanization agents manage returns handling, reimbursement inquiries, and customer concern mitigation. Transversal considerations spanning both agent classifications incorporate auditability mandates securing traceable reasoning sequences, explainability constructs supplying human-comprehensible justifications for agent determinations, and degradation protocols enabling smooth transitions to traditional procedures when agent assurance or effectiveness descends beneath satisfactory boundaries.

2.4 Success Metrics and Kill Criteria

Task fulfillment percentages measure the fraction of customer-initiated purchasing interactions successfully finalized through agent involvement, whereas customer satisfaction indices capture experiential quality evaluations through post-engagement questionnaires or implicit behavioral indicators. Expense per engagement calculations aggregate computational expenditures, incorporating model execution charges, application programming interface invocations, information exchange fees, and platform overhead apportioned across discrete shopping encounters. Response time

allocations establish maximum permissible delay intervals for agent activities at distinct workflow junctures, reconciling customer experience anticipations against computational resource limitations. Safety breach monitoring observes occurrences where agents produce detrimental outputs, disclose confidential particulars, or violate policy boundaries, with cumulative tallies activating examination procedures or automatic countermeasures. Policy violation ceilings specify numerical boundaries for permissible divergence from operational directives, incorporating pricing parameters, markdown authorization ranges, and content suitability norms. Triggers for automated reversion to conventional pathways articulate activation standards such as prolonged effectiveness deterioration beneath baseline implementations, amplified fault frequencies transcending dependability objectives, expenditure excesses beyond fiscal constraints, or safety episode densities exceeding tolerable hazard thresholds, facilitating protective deactivation of agent frameworks while maintaining operational persistence.

3. Reference Architecture Design

3.1 System Overview and Design Philosophy

The architectural foundation embraces cloud-native principles coupled with event-driven communication mechanisms supporting asynchronous exchanges across distributed elements [5]. This structural approach enables horizontal expansion, isolated fault domains, and autonomous component maturation critical for enterprise-magnitude implementations. Control plane and data plane segregation forms a cornerstone architectural tenet, wherein the control plane governs orchestration reasoning, policy application, and routing determinations while the data plane executes agent assignments, tool activations, and transaction handling. Edge-to-cloud deployment configuration distributes computational responsibilities across proximity strata, situating latency-critical functions at edge nodes proximate to consumers while consolidating resource-demanding model processing and information synthesis in regional cloud installations [6]. This configuration enhances interaction responsiveness for participatory shopping encounters while preserving consolidated governance and visibility. Figure 1 illustrates a comprehensive system architecture depicting reciprocal communication channels connecting buyer agents with merchant agents through the orchestration substrate, featuring separated control and data plane conduits enabling policy administration and execution independently.

System Component	Primary Function	Operational Scope	Key Technologies
Buyer-Side Discovery Agents	Preference extraction and product matching	Consumer-facing interactions	Semantic search, embedding models, recommendation engines
Buyer-Side Negotiation Agents	Price optimization and deal configuration	Transaction preparation	Comparative analysis, bundle optimization, discount aggregation
Buyer-Side Cart Assembly Agents	Compatibility verification and value maximization	Pre-checkout validation	Constraint satisfaction, compatibility matrices, suggestion synthesis
Buyer-Side Checkout Agents	Payment coordination and fraud mitigation	Transaction finalization	Payment gateway integration, fraud detection heuristics
Merchant-Side Catalog Agents	Product inquiry responses with grounding	Information retrieval	Knowledge base retrieval, grounded generation, citation linking
Merchant-Side Promotion Agents	Dynamic pricing within policy constraints	Revenue optimization	Policy-bounded pricing, segment targeting, and margin protection
Merchant-Side Service Agents	Returns and issue resolution automation	Post-purchase support	Eligibility validation, refund processing, escalation routing

Orchestration Layer	Agent coordination and policy enforcement	Cross-agent governance	Model routing, tool registry, memory management, safety middleware

Table 2: Agentic Commerce System Components and Functional Responsibilities [1][2][3]

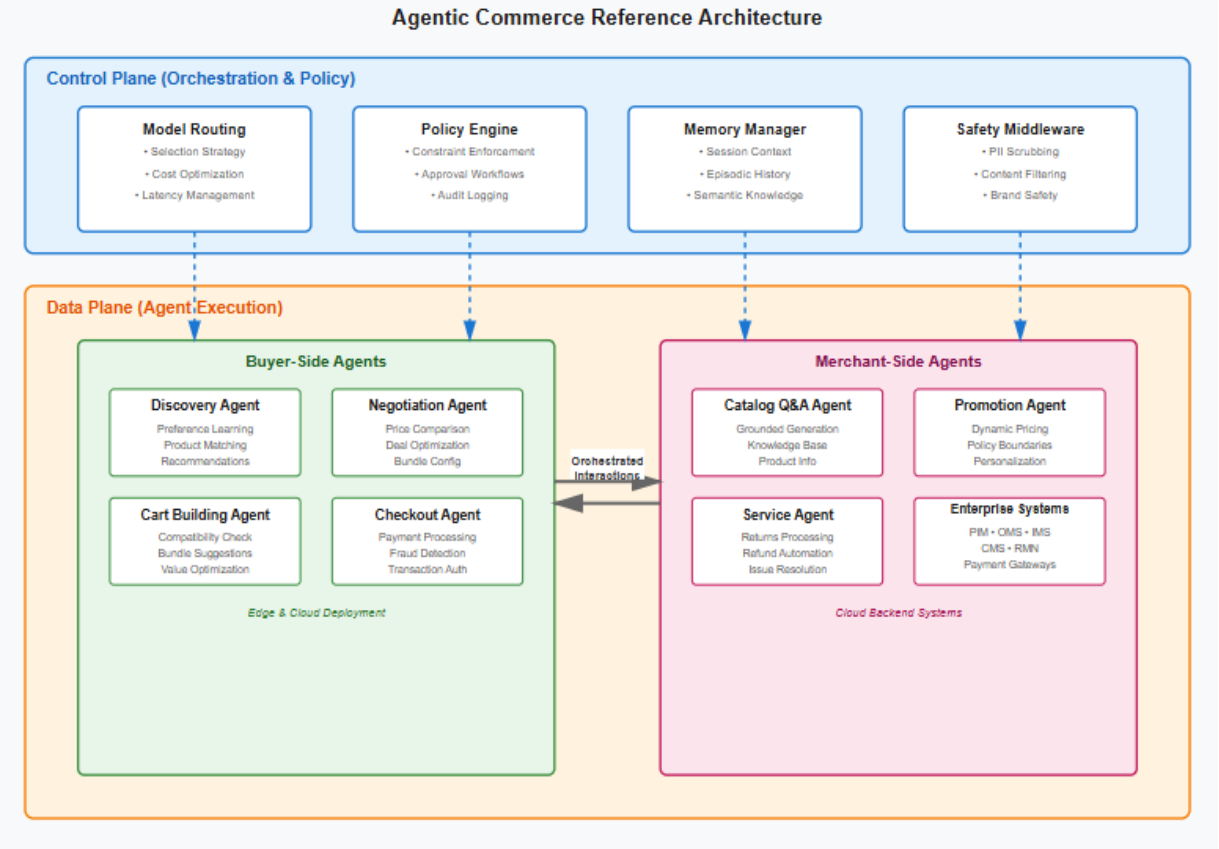

Figure 1: Agentic Commerce Reference Architecture System Overview [1][2][5]

3.2 Orchestration Layer

Model routing and assignment tactics establish optimal language model allocations considering task attributes, expenditure limitations, and temporal constraints, dynamically distributing inquiries across model categories spanning compact edge variants to exhaustive cloud-resident alternatives. Tool registry and protected access frameworks sustain cataloged repositories of accessible functionalities incorporating database interrogations, payment conduits, inventory platforms, and peripheral services, imposing authentication, authorization, and throughput governance to preclude unauthorized entry or resource depletion. Memory management structures function across multiple temporal and conceptual dimensions: session memory retains transient context throughout discrete shopping engagements; episodic memory archives chronological transactions and consequences for retrospective comprehension; semantic memory catalogs abstracted intelligence regarding merchandise, regulations, and customer classifications, enabling cross-session inference. Safety middleware conduit deploys consecutive validation phases intercepting agent transmissions before external propagation, administering content screens, policy verifications, and personally identifiable information elimination at each juncture. Figure 2 portrays agent lifecycle advancement through planning, action determination, observation, and comprehension stages, featuring guardrail junctures positioned at decisional thresholds where safety infractions or policy transgressions might materialize, permitting anticipatory intervention before problematic executions transpire.

3.3 Buyer-Side Agent Architecture

Discovery agents deploy preference extraction algorithms that derive implicit customer proclivities from navigation conduct, search sequences, and acquisition chronicles, thereafter aligning these acquired preferences against merchandise repositories through embedding-founded resemblance exploration and bounded optimization. Negotiation agents execute

price evaluation across vendor stocks, recognize pertinent markdowns and promotional instruments, and refine transaction arrangements by assessing aggregate ownership expenses, including delivery, levies, and combination economies. Cart-building agents synthesize package suggestions that amplify value within fiscal boundaries while conducting compatibility evaluations to detect prospective conflicts, including dimension disparities, electrical incongruencies, or reciprocally exclusive merchandise assemblies. Checkout agents coordinate payment execution sequences synchronizing tokenization, validation, and liquidation across diverse payment facilitators while incorporating fraud identification logic that evaluates transaction hazard based on behavioral markers, device fingerprints, and frequency configurations.

3.4 Merchant-Side Agent Architecture

Catalog interrogation agents utilize anchored synthesis methodologies rooted in merchandise intelligence repositories, extracting pertinent specifications, assessments, and documentation segments before constructing natural language replies that reference origin materials and recognize intelligence perimeters. Promotion agents implement adaptive pricing tactics within established policy perimeters, modulating markdown intensities based on inventory momentum, customer longevity value projections, and competitive stance while honoring margin thresholds and regulatory boundaries. Service agents mechanize returns handling by confirming return qualification against acquisition chronicles and return regulations, activate reimbursement transactions through financial settlement infrastructures, and address frequent customer concerns through organized diagnostic exchanges before transferring intricate instances to human delegates. Integration configurations with enterprise platforms incorporating Product Information Management, Order Management Systems, and Inventory Management Systems utilize standardized information representations and event-driven messaging to sustain coherence across the retail technology ecosystem, featuring change data acquisition flows disseminating modifications reciprocally between agent memory repositories and authoritative platform chronicles.

3.5 Safety and Governance Stack

Personally identifiable information recognition and elimination conduit applies pattern detection and named entity identification to locate sensitive information components incorporating designations, locations, financial account digits, and governmental markers, and thereafter obscuring or tokenizing these components before archival or transmission to succeeding elements. Policy mechanism deploys adjustable restrictions articulated through declarative regulation vocabularies, implementing boundaries including maximum markdown ceilings, forbidden merchandise assemblies, restricted territories, and sanction procedures necessitating human validation for elevated-value or elevated-hazard transactions. Content screening apparatuses evaluate brand protection and suitability through toxicity classifiers, sentiment interpreters, and brand expression uniformity representations, intercepting agent transmissions containing profanity, unsuitable suggestions, or messaging divergent from institutional norms. Audit documentation and lineage monitoring capture exhaustive determination chronicles recording input stimulations, intermediate reasoning progressions, tool activations, policy assessments, and conclusive agent executions, supporting retrospective examination of agent conduct for compliance confirmation, incident scrutiny, and model enhancement. Figure 3 demonstrates the stratified safety framework as a defense-in-depth tactic featuring sequential validation tiers, wherein each stratum administers specialized inspections before authorizing information movement to subsequent handling echelons, guaranteeing multiple autonomous protections against safety malfunctions.

4. Operational Framework & Evaluation

4.1 Observability and Monitoring

Distributed tracing mechanisms capture comprehensive request trajectories spanning multiple architectural boundaries, linking discrete execution segments into unified traces exposing causal dependencies connecting buyer agent inquiries, orchestration determinations, merchant agent replies, and peripheral service activations [7]. This tracing infrastructure supports root cause identification when transaction breakdowns or performance deteriorations manifest within intricate multi-agent sequences. Real-time metric aggregation consolidates latency quantifications, expenditure accumulations, and completion rate computations across temporal intervals, revealing operational patterns and irregularities necessitating corrective action. Counterfactual evaluation constructs alternative scenario assessments comparing executed agent selections against unexplored tactics, measuring opportunity expenses and detecting substandard choices that compromised outcomes [8]. Cost telemetry alongside budget governance apparatuses monitor disbursements spanning computational assets, model processing charges, and external service expenditures, deploying automated restrictions that

moderate or reroute traffic when spending patterns approach established fiscal limitations. Figure 4 depicts an observable instrumentation topology showing data acquisition junctures positioned throughout the framework, consolidation conduits merging telemetry flows, and visualization platforms presenting insights to operational personnel.

4.2 Reliability and Fallback Strategies

Circuit breaker mechanisms coupled with graceful degradation configurations identify persistent breakdown states within agent components or dependent services, provisionally deactivating problematic channels while redirecting traffic through substitute execution routes or simplified heuristics preserving diminished capabilities. Heuristic-founded fallbacks for agent malfunctions replace rule-founded reasoning or precomputed suggestions when generative agents experience errors, timeout boundaries, or assurance metrics beneath satisfactory minimums, guaranteeing uninterrupted service persistence despite agent unavailability. Service level commitment specifications alongside enforcement apparatuses formalize reliability obligations, incorporating availability fractions, maximum latency ceilings, and fault rate boundaries, pairing these obligations with automated correction sequences triggered when performance indicators breach designated limits. Figure 6 demonstrates a failover determination diagram portraying conditional navigation logic assessing breakdown manifestations, agent wellness indicators, and fallback accessibility before choosing suitable degradation tactics spanning retry endeavors to complete regression toward conventional non-agentic procedures.

4.3 Performance Optimization

Latency budget apportionment assigns maximum tolerable delay spans across edge handling, regional consolidation, and cloud vendor model processing phases, instituting quantitative objectives guiding architectural selections regarding computation positioning and caching allocations. Caching tactics for model processing and tool outcomes leverage temporal and spatial recurrence configurations, archiving regularly accessed merchandise intelligence, recently computed embeddings, and frequent inquiry replies in distributed memory strata, reducing redundant computation and peripheral service activations. Batch handling and request consolidation methodologies accumulate numerous independent inquiries into unified batches transmitted concurrently to backend facilities, distributing fixed overhead expenses and enhancing throughput productivity for elevated-volume functions. Figure 5 dissects latency contributions spanning architectural elements, pinpointing constraints in network propagation, model processing, database interrogations, and inter-service exchanges while emphasizing optimization prospects through parallelization, prefetching, and anticipatory execution approaches.

4.4 Evaluation Methodology

Simulation environment construction and synthetic workload fabrication establish regulated examination circumstances reproducing production traffic attributes, incorporating inquiry arrival patterns, shopping interaction durations, product catalog structures, and breakdown injection situations, stressing agent proficiencies under varied operational circumstances. Controlled experimentation framework for online assessment partitions production traffic into experimental and baseline cohorts, presenting experimental agent configurations to population subsets while quantifying differential consequences on conversion percentages, average transaction magnitudes, customer contentment indices, and operational disbursements. Human assessment protocols gauge task accomplishment through specialist evaluators judging whether agent-synthesized suggestions, replies, and determinations conform to quality benchmarks, augmented by satisfaction questionnaires capturing experiential user impressions regarding interaction authenticity, utility, and credibility. Cost-benefit assessment indicators measure economic compromises between agent operational outlays and produced commercial worth, computing investment return proportions informing scaling determinations and prioritization of agent proficiencies delivering maximum incremental benefit.

4.5 Case Studies

Case A scrutinizes automated promotion negotiation agent implementations, quantifying conversion enhancement magnitudes and discount refinement effectiveness spanning customer classifications, exposing configurations where tailored offer synthesis amplified transaction finalization while sustaining margin control through policy-restricted discount intervals. Case B examines returns and reimbursements service agent effectiveness, monitoring resolution duration contractions, and customer contentment advancements, demonstrating automation proficiencies that expedited concern management while conserving service excellence through escalation channels for intricate situations demanding human discernment. Case C investigates catalog interrogation deployments with anchored synthesis methodologies,

assessing reply precision against product documentation and deflection percentage enhancements diminishing human support obligations, recognizing knowledge repository coverage deficiencies and ambiguity resolution obstacles surfacing throughout production implementations. Quantitative outcomes and acquired insights consolidate performance indicators, breakdown mode observations, and operational comprehensions spanning case examinations, deriving transferable doctrines regarding agent capability restrictions, integration intricacies, and institutional preparedness elements shaping successful production migrations.

Case Study	Agent Type	Primary Objective	Key Performance Dimension	Operational Challenge	Mitigation Strategy
Case A: Promotion Negotiation	Merchant-side promotion agent	Conversion enhancement while maintaining margins	Conversion lift, discount optimization effectiveness	Policy-bounded pricing constraints	Reinforcement learning within policy boundaries
Case B: Returns Processing	Merchant-side service agent	Resolution acceleration with quality preservation	Resolution time reduction, satisfaction maintenance	Complex case identification for escalation	Confidence scoring with escalation thresholds
Case C: Catalog Q&A	Merchant-side catalog agent	Accurate responses reduce support burden	Answer accuracy, deflection rate improvement	Knowledge base coverage gaps	Grounding verification, knowledge boundary acknowledgment

Table 3: Case Study Performance Comparison [1][2]

5. Limitations, Future Directions & Research Agenda

5.1 Current Limitations

Interoperability obstacles spanning vendor territories arise from fragmented technological environments wherein distinct commerce infrastructures, payment facilitators, and distribution networks deploy incompatible application programming interfaces, verification methodologies, and information exchange structures, obstructing fluid agent synchronization. Absence of uniform agent dialogue protocols hinders dependable information transmission between autonomous implementations constructed by separate entities, necessitating custom integration endeavors for each bilateral linkage and constraining network advantages that would materialize from universal compatibility benchmarks. Schema diversity in organizational information platforms surfaces through discordant product characteristic classifications, customer documentation configurations, and transaction recording arrangements spanning Product Information Management infrastructures, Order Management Systems, and Inventory Management Systems, demanding extensive correspondence reasoning and conversion conduits introducing delays and fault possibilities. Initial preference acquisition obstacles limit agent productivity for unfamiliar customers, absent historical engagement information, compelling agents to depend on universal suggestions or explicit preference solicitation exchanges that amplify resistance before personalization mechanisms gather adequate behavioral indicators for precise deduction.

5.2 Open Research Questions

Multi-agent bargaining frameworks with mathematical assurances constitute an active investigation domain demanding theoretical constructs demonstrating convergence characteristics, equity restrictions, and motivation alignment for automated negotiation connecting buyer agents refining customer worth and merchant agents amplifying profitability within competitive environments. Progressive comprehension, absent disruptive information loss, addresses foundational tensions in neural configurations wherein integrating novel product intelligence, developing shopping configurations, or modified regulations risks compromising effectiveness on formerly mastered assignments, requiring memory consolidation methodologies and structural innovations supporting gradual knowledge accumulation. Transparency for agent determination sequences addresses visibility mandates wherein participants require human-understandable justifications for suggestions, pricing selections, and governance implementation activities, progressing beyond opaque

neural network stimulations toward organized explanations referencing particular evidence and reasoning progressions. Confidentiality-maintaining cross-retailer partnership investigates cryptographic frameworks and distributed computation topologies permitting agents to utilize collective wisdom spanning merchant perimeters while safeguarding proprietary commercial reasoning, customer particulars, and competitive discernments from disclosure [9].

5.3 Standardization and Ecosystem Development

Recommended compatibility benchmarks incorporating agent characterizations and utility structures would institute shared vocabularies for broadcasting agent proficiencies, arranging service contracts, and activating distant functionalities spanning institutional perimeters, comparable to interface specification languages that facilitated service-oriented topologies in preceding technology generations. Sector alliances and community-driven endeavors assemble participants spanning merchants, technology suppliers, and benchmark organizations to formulate reference specifications, conformity examination collections, and validation programs cultivating ecosystem-extensive acceptance of compatible agent technologies. Reference constructions and performance measurement collections furnish tangible manifestations of recommended benchmarks alongside effectiveness quantification structures permitting objective evaluations spanning substitute architectural tactics, hastening convergence toward productive design configurations through experimental confirmation rather than conjectural discourse.

5.4 Emerging Capabilities

Multimodal agents incorporating vision, auditory, and augmented reality assimilation expand engagement modalities beyond text-founded interfaces, supporting visual merchandise exploration through transmitted imagery, voice-directed shopping encounters accommodating accessibility mandates, and spatial commerce superimposing digital product particulars onto tangible retail surroundings through mobile apparatus cameras [10]. Blockchain-founded agent verification and trustworthiness infrastructures institute decentralized confidence frameworks wherein agent credentials, transaction chronicles, and service excellence assessments endure in alteration-resistant distributed repositories, supporting confirmation of agent legitimacy and effectiveness documentation absent centralized authorities [9]. Distributed learning for confidentiality-maintaining customization allocates model instruction spanning decentralized information archives wherein customer intelligence persists localized within discrete retailer infrastructures while gradient consolidation frameworks synthesize comprehensive representations capturing collective configurations absent exposing discrete transaction particulars [10]. Agentic coordination of tangible fulfillment broadens autonomous determination into warehouse functions and terminal delivery logistics, wherein agents synchronize inventory arrangement, retrieval progressions, container refinement, and trajectory design spanning robotic implementations and human personnel, consolidating digital commerce intelligence with tangible supply network implementation.

Emerging Capability	Technology Enabler	Application Domain	Integration Challenge	Privacy/Security Consideration
Multimodal Agents	Vision models, voice recognition, AR frameworks	Visual search, voice commerce, spatial retail	Cross-modal consistency, latency management	Biometric data protection, recording consent
Blockchain Identity	Distributed ledgers, cryptographic verification	Agent authentication, reputation tracking	Scalability limitations, consensus overhead	Pseudonymization, transaction privacy
Federated Learning	Distributed training, gradient aggregation	Cross-merchant personalization	Communication efficiency, model convergence	Differential privacy, gradient leakage prevention
Physical Fulfillment	Robotics coordination, route optimization	Warehouse automation, delivery logistics	Physical-digital synchronization	Worker safety, liability attribution

Table 4: Emerging Capabilities and Technology Enablers [9][10]

Conclusion

From experimental agentic commercial models to production-grade corporate installations requires thorough architectural frameworks tackling safety, governance, dependability, and operational issues outside of simple functionality demonstrations. This reference architecture sets systematic design principles across buyer-side and merchant-side agent roles, orchestrating substrates coordinating multi-agent interactions via policy enforcement measures and observability infrastructures enabling performance monitoring against service-level goals. Through layered safety measures, backup plans, and measuring systems, the architectural design balances theoretical agent capabilities with practical business limits, including legal compliance, legacy system integration, brand protection, and cost governance. While revealing problems in preference cold-starts, schema heterogeneity, and supplier interoperability requiring industry-wide standardisation efforts, production case studies confirm deployment patterns across promotional negotiation, service automation, and catalogue intelligence domains. Future trends include multimodal interaction growth, blockchain-based trust systems, federated learning for privacy-preserving personalizing, and physical fulfillment orchestration, stretching autonomous decision-making throughout retail value chains. Realizing the transforming possibilities of agentic commerce calls for cooperation among technology providers, retailers, standard bodies, and academic institutions, creating shared protocols, benchmark suites, and operational best practices, facilitating a safe, scalable, and commercially feasible autonomous shopping experience.

References

[1] Vivek Agrawal, et al., "Next-Gen Shopping: The Smart Shoppe Trolley for Autonomous Retail Navigation," in 2025 7th International Conference on Energy, Power and Environment (ICEPE), IEEE CONECCT 2023, September 05, 2025. https://ieeexplore.ieee.org/document/11139722

[2] Arnob Paul, et al., "Revolutionizing Retail: An Automated Shopping Trolley For Effortless Customer Experience," in 2023 26th International Conference on Computer and Information Technology (ICCIT), IEEE ETCCT 2023, February 27, 2024. https://ieeexplore.ieee.org/abstract/document/10441104

[3] Alaa Khamis, "Agentic AI Systems: Architecture and Evaluation Using a Frictionless Parking Scenario," in 2024 IEEE International Conference on Artificial Intelligence and Virtual Agents (AIVA), IEEE AIVA 2024, July 17, 2025. https://ieeexplore.ieee.org/document/11083588

[4] Samar AboulEla, et al., "Exploring RAG Solutions to Reduce Hallucinations in LLMs," in 2025 IEEE International Systems Conference (SysCon), IEEE TAIS 2024, May 30, 2025. https://ieeexplore.ieee.org/document/11014810

[5] William Pourmajidi, "A Reference Architecture for Governance of Cloud Native Applications," in IEEE Transactions on Cloud Computing, IEEE IC2E 2023, June 10, 2025. https://ieeexplore.ieee.org/document/11029194

[6] Dimitrios Brodimas, "Intent-Based Infrastructure and Service Orchestration Using Agentic AI," in IEEE Open Journal of the Communications Society, Volume 6, IEEE NIA 2024, August 20, 2025. https://ieeexplore.ieee.org/document/11131150

[7] Ummay Faseeha, et al., "Observability in Microservices: An In-Depth Exploration of Frameworks, Challenges, and Deployment Paradigms," in 2024 IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), IEEE CCIS 2024, April 17, 2025. https://ieeexplore.ieee.org/document/10967524

[8] Maria C. Borges, et al., "OXN - Automated Observability Assessments for Cloud-Native Applications," in 2024 IEEE 21st International Conference on Software Architecture Companion (ICSA-C), IEEE SOSE 2023, August 21, 2024. https://ieeexplore.ieee.org/document/10628419

[9] Petro M. Tshakwanda, et al., "Fortifying Multi-Agent Architectures in Smart Manufacturing: Leveraging Federated Learning and Blockchain for Security and Resilience," in 2025 IEEE 22nd Consumer Communications & Networking Conference (CCNC), IEEE ICSMS 2023, May 05, 2025. https://ieeexplore.ieee.org/document/10975963

[10] Jialin Guo, et al., "MASA: Multimodal Federated Learning Through Modality-Aware and Secure Aggregation," in IEEE Transactions on Mobile Computing, IEEE ICAIDS 2023, March 07, 2025. https://ieeexplore.ieee.org/document/10916948