

Retail Edge: Scalable POS System with Receipt Digitization for the Unorganized Sector

Amatya Joshi
Dept. of Computer Science and
Engineering
MIT ADT University
Pune, India
joshiamatya@gmail.com

Meet Acharya
Dept. of Computer Science and
Engineering
MIT ADT University
Pune, India
mastermeet24@gmail.com

Prof. Priyanka Gaikwad
Dept. of Computer Science and
Engineering
MIT ADT University
Pune, India
priyanka.gaikwad@mituniversity.edu.in

Abstract—RetailEdge is a scalable POS system that enhances transaction processing and inventory management. It features a receipt tracker that allows retailers to upload PDF receipts, which are digitized using OCR and automatically updated into the system's database. The system provides real-time processing and cloud-based analytics, improving operational efficiency and decision-making.

Keywords—Point-of-Sale (POS) System, Receipt Digitization, Optical Character Recognition (OCR), Retail Management, Data Automation, Inventory Management, Cloud Analytics, Transaction Processing, Scalable Systems, Retail Efficiency

I. INTRODUCTION

India's unorganised retail sector—comprising small kirana stores, street vendors, and independent shopkeepers—represents the majority of the country's retail market, yet remains underserved by digital technology. These retailers often rely on manual billing, paper receipts, and fragmented inventory systems, which create inefficiencies, hinder data management, and limit long-term scalability. The absence of affordable, intuitive digital solutions further exacerbates this divide, leaving these businesses excluded from the growing wave of data-driven retail transformation.

RetailEdge addresses this challenge by offering a lightweight, scalable Point-of-Sale (POS) system tailored for the operational and infrastructural constraints of the unorganised retail ecosystem. It supports real-time billing, inventory tracking, customer data handling, and works efficiently in both online and offline environments. The system's modular backend-agnostic architecture enables flexible deployment across local servers and cloud platforms, while maintaining resilience through features like asynchronous data sync and role-based access control.

A standout feature of RetailEdge is its integrated Receipt Tracker, which allows retailers to digitize physical receipts via scanned PDFs using Optical Character Recognition (OCR). The system extracts structured data—such as item names, prices, taxes, and timestamps—and updates the digital ledger automatically. This enables small retailers to retroactively build and maintain electronic sales records, paving the way for better inventory control, analytics, and regulatory compliance without disrupting existing workflows.

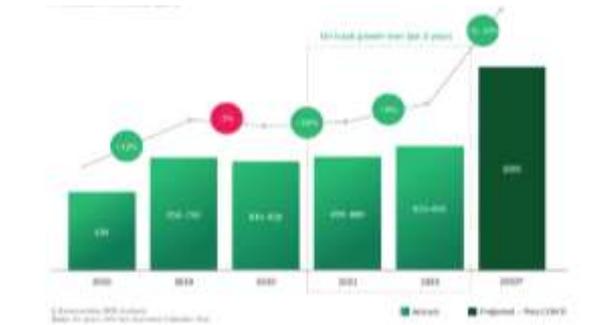


Fig. 1. India's Retail Market Size

A. Motivation

India's unorganised retail sector comprises over 12 million small and micro businesses that form the backbone of the country's consumer economy. Despite their economic significance, these retailers often operate without access to modern digital tools due to cost barriers, infrastructure limitations, and lack of technical expertise. Daily operations such as sales tracking, inventory updates, and customer record maintenance are typically handled through manual, paper-based systems—prone to errors, inefficiencies, and data loss. This lack of digitization not only hampers operational visibility but also restricts their ability to scale, comply with regulatory norms, and leverage government schemes aimed at digital empowerment.

Existing POS systems are largely designed with organized retail in mind and assume the presence of stable internet connectivity, sufficient computing infrastructure, and skilled operators. Furthermore, legacy data—especially in the form of physical receipts—remains unutilized in transitioning businesses, representing a bottleneck in digital onboarding. The absence of a unified solution that bridges daily retail operations with automated digitization of past transactions has created a critical gap.

RetailEdge is conceived as a response to this need—motivated by the goal of delivering an affordable, low-footprint POS system that includes an OCR-based module for retroactive receipt digitization. This approach empowers retailers not only to manage ongoing transactions efficiently but also to transform historical paper-based data into a usable digital format. By addressing both current workflow management and data backlog challenges, RetailEdge contributes to long-term

operational resilience and digital maturity in the informal retail space.

B. Objective

This research aims to design and implement a robust Point-of-Sale (POS) system tailored to the infrastructural and operational constraints of India's unorganised retail sector. The primary goal is to empower small retailers with a digital solution that is both scalable and resource-efficient. The key objectives include:

- **Real-Time Transaction Processing:** Enable efficient handling of sales, inventory, and billing activities through a lightweight, intuitive POS interface with support for real-time updates.
- **Receipt Digitization via OCR:** Integrate an Optical Character Recognition (OCR) module that extracts structured data from scanned PDF receipts, facilitating the digitization of legacy records into the retailer's system database.
- **Infrastructure Agnosticism:** Ensure seamless deployment across both local servers and cloud platforms, with offline operability, asynchronous syncing, and minimal hardware dependency.
- **Modular System Architecture:** Develop a component-based design that supports extensibility for analytics, compliance tools, and third-party API integrations, without affecting core system stability.
- **Data Accessibility and Insight:** Provide basic reporting and visual dashboards to support better decision-making, inventory planning, and historical sales tracking for non-technical users.

C. Scope

RetailEdge is focused on the digital enablement of unorganised retail operations through lightweight POS infrastructure and document digitization. The current research scope includes: real-time sales processing and inventory tracking; offline-compatible data storage and synchronization; OCR-based extraction of transactional data from uploaded receipt PDFs; and structured integration of this data into the system database. The system will expose modular REST APIs to support transaction logging, inventory updates, and receipt ingestion workflows.

The architecture is designed to be backend-agnostic and deployable across both cloud-based and on-premise environments, with low system overhead to suit resource-constrained devices. While core features emphasize usability and operational efficiency, advanced modules such as predictive analytics, multi-branch coordination, GST compliance automation, and integration with third-party retail intelligence platforms are considered beyond the current scope and marked for future work. Security measures include role-based access control and basic data validation; however, full-scale encryption and regulatory compliance features are excluded from the present implementation.

This research is primarily targeted at India's small-scale retailers but is extensible to similar informal markets globally, providing a foundational framework for scalable digital retail systems in developing economies.

II. RELATED WORK

A. Prior Systems

The digital transformation of retail operations has largely centered around organized retail chains and enterprise-level POS systems, with limited penetration into the unorganised sector. Several commercial solutions such as **Tally ERP**, **Zoho POS**, and **Shopify POS** offer comprehensive billing and inventory features, yet they often come with high licensing costs, steep learning curves, or infrastructure dependencies unsuitable for small retailers operating in resource-constrained settings.

Open-source platforms like Odoo and Floreant POS have attempted to bridge this gap by offering modular architectures and community-driven support. However, these platforms typically lack native support for digitizing legacy transactional data, such as paper-based receipts, which continue to dominate record-keeping in the unorganised sector. Additionally, most existing systems rely on consistent internet connectivity and assume a certain baseline of digital literacy, further limiting adoption.

While research efforts have been made in OCR and document processing domains, they are generally

Feature	RetailEdge	Shopify POS	Square POS	Zoho Inventory	Vyapar
Built for Indian Unorganized Retail	✓	X	X	X	✓
GST Billing + RazorPay Support	✓	X	X	✓	✓
Offline Functionality	✓	X	✓	X	✓
High Customization & Modern Stack	✓	X	X	✓	X
AI Chatbot + Analytics Dashboard	✓	X	X	✓	X

Table. 1. Feature-Based Comparison of RetailEdge with Major Retail Software Solutions

isolated from retail-oriented workflows. Studies and tools such as Tesseract OCR and Google Cloud Vision offer robust recognition capabilities, yet their integration into POS systems remains minimal, especially in the context of low-cost, offline-capable environments.

There remains a clear research and implementation gap in providing an end-to-end POS ecosystem that is cost-effective, offline-resilient, receipt-digitization-enabled, and specifically tailored for the informal retail landscape. RetailEdge addresses this gap by combining transaction management with OCR-powered data migration in a form factor accessible to India’s unorganised retail community.

B. Literature Review

Patel *et al.* [1] explored the adoption of modular POS systems by small retailers, highlighting that ease of use, low-cost deployment, and offline capabilities are essential features for informal retail environments. However, their research also noted that mainstream solutions such as Shopify POS and Zoho are often not viable for small retailers due to financial and infrastructural limitations.

Gupta and Sharma [2] further emphasized these challenges, particularly within the Indian unorganised retail sector. Their study revealed that complex licensing models, lack of regional language support, and internet dependency are key barriers preventing widespread POS adoption in rural and semi-urban markets.

Rao and Srinivasan [3] analyzed open-source POS systems like Odoo and Floreant POS. While these systems provide higher flexibility and customization, they often lack integration with legacy data systems, making them less

effective for retailers who want to migrate from paper-based receipts to digital systems.

On the digitization front, Smith [4] introduced the Tesseract OCR engine, which became a foundational tool in open-source document processing. Li *et al.* [5] built on this work, demonstrating how pre-processing methods like image binarization and noise reduction can significantly improve OCR accuracy, especially for structured documents such as receipts.

Mehta and Bansal [6] identified the pressing need for domain-specific OCR applications, particularly in retail, where past transactional data on paper must be digitized and synchronized with current databases. They concluded that combining POS functionality with a receipt digitization module would enhance operational efficiency and support digital transitions in small businesses.

III. PROPOSED SYSTEM

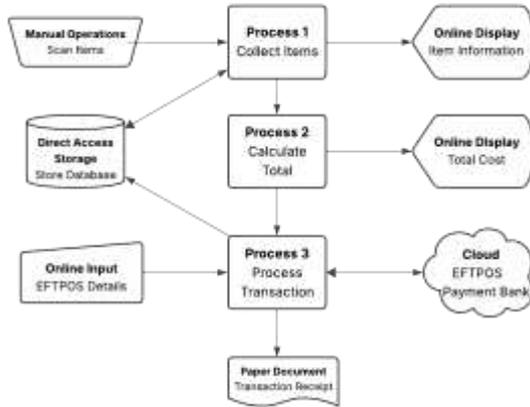


Fig. 2. System Process FlowChart

A. Conceptual Framework

RetailEdge is designed as a modular, scalable Point-of-Sale (POS) system that addresses the technological gaps faced by the unorganised retail sector in India. The core concept revolves around providing a lightweight, offline-capable POS infrastructure that is easy to deploy, affordable, and integrable with legacy retail practices. Recognizing the operational constraints of small vendors—such as inconsistent internet access, non-standardized inventory practices, and lack of digital literacy—the system is built to function with minimal setup while ensuring reliable transaction recording and reporting.

One of the key innovations in RetailEdge is the receipt digitization module, powered by Optical Character Recognition (OCR). This add-on allows users to upload scanned or photographed receipts (in PDF or image formats), from which structured data is extracted and stored in the POS database. This enables retailers to transition from paper records to digital logs without needing to re-enter historical data manually. The OCR engine is trained specifically on Indian retail receipt formats, allowing it to recognize item names, quantities, rates, GST, and date-time metadata with high accuracy.

The architecture emphasizes modularity and separation of concerns. The POS core manages transactions, inventory, billing, and reporting, while the OCR module acts as an auxiliary service, interfacing with the document processing layer. Data integrity is maintained through schema validation and confidence-based extraction scoring to flag uncertain fields. The system exposes RESTful APIs for integration with other retail tools, banking systems, and analytics dashboards, ensuring interoperability.

RetailEdge thereby functions as both a live transaction manager and a digitization engine for past records. By blending real-time billing with post-facto data ingestion from paper receipts, it bridges the digital divide in retail recordkeeping—empowering small retailers to organize operations, maintain compliance, and derive business insights from previously inaccessible data.

B. System Overview

RetailEdge is built around two key modules: the Point-of-Sale (POS) Core and the Receipt Digitization Engine. The POS Core manages core retail functions such as transaction processing, inventory tracking, GST-compliant billing, and customer management. Designed to run on a variety of hardware, including mobile and low-power devices, it supports offline operations with automatic data synchronization once connectivity is restored. A responsive, user-friendly interface enables even digitally inexperienced retailers to process sales, monitor stock, and generate real-time business reports efficiently.

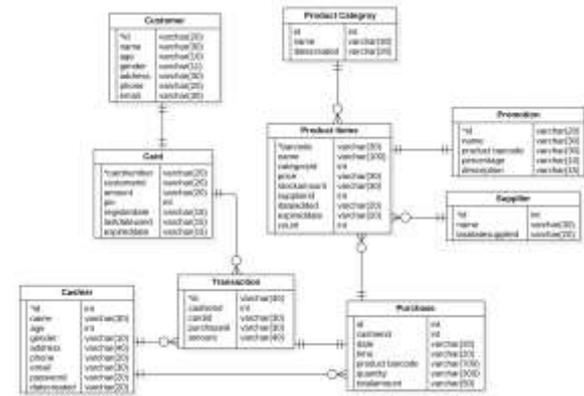


Fig. 3. System ER Diagram

Complementing the POS module is the Receipt Digitization Engine, which enables users to upload physical receipts in PDF or image format. A custom-trained OCR pipeline extracts structured data—item names, prices, taxes, and timestamps—into the system’s database. This aids in converting legacy physical records into searchable digital formats, streamlining bookkeeping and enabling analytics on past sales. The system features a unified dashboard for monitoring both live and uploaded data, API endpoints for integrations, and robust security features like RBAC and encrypted storage. RetailEdge can be deployed both on the cloud or locally, making it highly adaptable to the infrastructure conditions of the unorganised retail sector.

IV. SOFTWARE REQUIREMENT SPECIFICATION

A. Functional Requirements

- POS Operations: Real-time GST-compliant billing, inventory control, transaction logging, and invoice generation.
- User Authentication: Secure sign-in, session handling, and role-based access using Clerk.
- Receipt Digitization: Upload receipts (PDF/images), OCR-based data extraction, validation, and storage.
- Analytics & Reporting: Generate visual sales reports, transaction summaries, and business insights.
- AI Routing & Processing: Use Inngest and AgentKit to manage background receipt processing tasks and AI-based field mapping.

- **Payment Gateway Integration:** Stripe and RazorPay-powered digital payment support for a wide range of payment methods.
 - **Feature Management:** Schematic is used to manage feature rollouts and optimize user experience.
- B. Non-Functional Requirements*
- **Platform Agnostic:** Fully web-based system powered by Next.js for seamless performance across devices.
 - **Scalability:** Modular design allows extension of features, such as loyalty programs or multilingual support.
 - **Security:** OAuth-based auth via Clerk, secure session tokens, HTTPS communication, and role-based access control.
 - **Performance & Reliability:** Serverless hosting via Vercel for rapid deployments and auto-scaling; robust AI task queue handling via Inngest.
 - **User Experience:** Clean, responsive UI built with ShadCN and Tailwind CSS ensures accessibility for non-tech-savvy users.
 - **Data Accuracy:** Convex and MySQL databases ensure structured and reliable data storage across modules.

C. Software Stack

a) POS System

- **Frontend & Backend:** Next.js, React.js, Redux, TypeScript
- **Database:** MySQL
- **Authentication:** Clerk
- **UI & Styling:** ShadCN, Tailwind CSS
- **Hosting:** Vercel
- **Payment Gateway:** Stripe, RazorPay

b) Receipt Tracker

- **Frontend & Backend:** Next.js, React.js, Redux, TypeScript
- **Database:** Convex
- **AI & Task Routing:** Inngest, AgentKit
- **Authentication:** Clerk
- **UX & Feature Management:** Schematic
- **Payment Gateway:** Stripe, RazorPay
- **Hosting:** Vercel

V.METHODOLOGY

The development of RetailEdge follows a modular, scalable, and efficient approach, leveraging state-of-the-art technologies to address the needs of small and medium-sized retail businesses. The methodology encompasses system architecture, development processes, integration, and deployment. It focuses on delivering an easy-to-use yet robust solution, with specific emphasis on both the POS System and Receipt Tracker.

A. System Architecture

RetailEdge is built using a Microservices architecture, with two key modules: the POS System and the Receipt Tracker.

Both modules operate independently but share common services for authentication and payment processing.

a) POS System

- **Core Functions:** Sales transactions, inventory management, GST-compliant billing, and invoice generation.
- **Frontend Technology:** Developed using Next.js, React.js, and Redux for responsive, state-managed user interactions.
- **Backend Database:** Uses MySQL to store transaction data and inventory updates.
- **AI Integration:** Predictive analytics for demand forecasting and AI-powered chatbot for user assistance.
- **Authentication:** Handled through Clerk for secure login and role-based access control.

b) Receipt Tracker

- **Core Functions:** Allows retailers to upload physical receipts (PDF/images), digitizes them using OCR technology, and stores the extracted data.
- **Frontend Technology:** Built with Next.js and React.js to ensure a seamless user experience.
- **Backend Database:** Uses Convex, a cloud-native database, for real-time data storage and access.
- **AI Integration:** Inngest and AgentKit for managing receipt processing tasks and field mapping.

B. Development Process

a) Requirement Gathering & Analysis

- **Objective:** Understand the needs of small and medium-sized retailers, focusing on affordability, scalability, and ease of use.
- **Key Features Identified:** Real-time sales tracking, GST-compliant billing, receipt digitization, predictive analytics, and secure authentication.

b) Design & Prototyping

- **UI/UX Design:** Designed using ShadCN and Tailwind CSS to ensure a clean, minimalistic, and user-friendly interface.
- **Prototyping:** Wireframes and prototypes developed to visualize the system's interaction flow.

c) Implementation

- **Frontend Development:** The frontend is built using Next.js, React.js, and Redux to ensure optimal user interaction and smooth state management.
- **Backend Development:** Developed using Next.js for server-side rendering (SSR), ensuring improved performance and SEO.
- **Payment Integration:** Integrated Stripe and RazorPay for secure payment processing, supporting a wide range of digital payment methods.
- **Authentication:** Clerk is used for secure user authentication and session management.

d) AI & OCR Integration

- **Receipt Digitization:** OCR-based algorithms are used to extract relevant data from uploaded receipts (PDF/images).

- AI-Powered Analytics: Integrated predictive analytics within the POS System for demand forecasting and a chatbot for general assistance.
- Background Task Management: Inngest handles background processing tasks, while AgentKit routes tasks efficiently for seamless user experience.

e) *Testing & Quality Assurance*

- Testing Frameworks: Automated testing is conducted using Jest (unit testing) and Cypress (integration and end-to-end testing).
- Testing Focus: The system is rigorously tested for functionality, security, and performance across both the POS System and Receipt Tracker modules.

f) *Deployment*

- Hosting Platform: The system is deployed on Vercel, providing serverless hosting that ensures auto-scaling, rapid deployments, and high performance.
- Monitoring & Maintenance: Vercel’s robust infrastructure ensures uptime and reliability, even during peak traffic.

- Build backend, frontend, and AI integrations.
- Deliverables: POS system (Next.js, MySQL), receipt tracker (OCR, Inngest), payment gateways (Stripe, RazorPay), AI and predictive analytics, development documentation.

Phase 4: Testing & Quality Assurance

- Perform unit, integration, and end-to-end testing.
- Deliverables: Bug fixing, UAT, final review in staging, test cases documentation.

Phase 5: Documentation

- Complete system, user manuals, and API documentation.
- Deliverables: Final documentation for users, admin, and developers.

Phase 6: Deployment & Go Live

- Deploy system to production and ensure operational readiness.
- Deliverables: Go-live deployment, post-launch support, deployment documentation.

VI. PROJECT PLAN

The RetailEdge project followed a structured timeline:

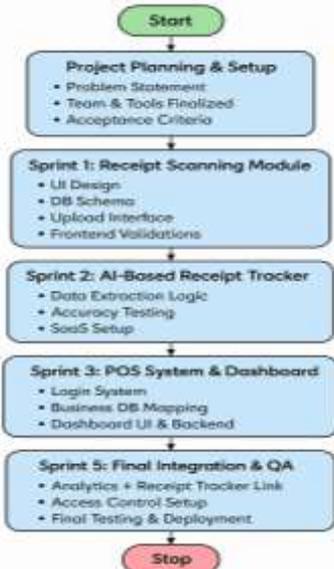


Fig. 4. Project Plan

Phase 1: Planning & Requirement Gathering

- Understand business requirements and finalize core features.
- Deliverables: Requirement documentation, architecture design, third-party integrations, initial documentation outline.

Phase 2: Design & Prototyping

- Develop wireframes and UI/UX prototypes.
- Deliverables: High-fidelity designs, user flow diagrams, feedback and iterations, initial design documentation.

Phase 3: Development & Implementation

VII. RESULTS

RetailEdge successfully delivers a robust, dual-module solution tailored for the unorganised retail sector. The POS System demonstrated seamless handling of real-time billing, inventory tracking, GST-compliant invoicing, and integrated payment processing through Stripe and RazorPay. AI-based predictive analytics provided actionable sales insights, while a built-in chatbot enhanced user assistance and engagement. The Receipt Tracker efficiently digitized receipts using OCR, with accurate data extraction and real-time processing managed by Inngest and AgentKit. The system’s intuitive interface, built with ShadCN and Tailwind CSS, ensured a smooth and accessible user experience across devices. Combined, these modules enhanced operational efficiency and simplified day-to-day retail management.

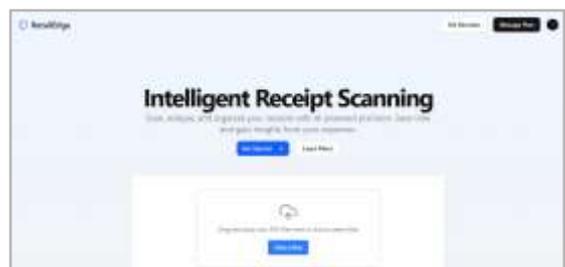


Fig. 5. Retail Edge Receipt Tracker Landing Page



Fig. 6. Retail Edge Receipt Tracker Authentication using Clerk



Fig. 7. Stripe Product Catalog

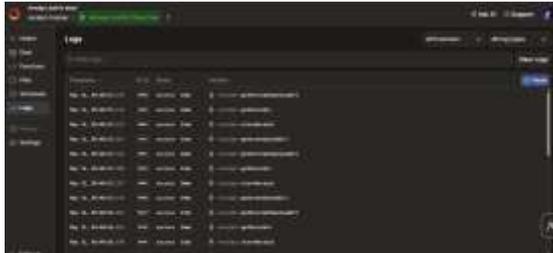


Fig. 8. Convex Database and Logging



Fig. 9. Inngest AI Agent Workflow



Fig. 10. RetailEdge Receipt Tracker Manage Plan



Fig. 11. Schematic Product Features

VIII.DISCUSSION

The modular architecture ensured system scalability, maintainability, and independent evolution of both the POS and Receipt Tracker modules. Leveraging modern frameworks like Next.js, React.js, and Redux enabled efficient rendering and state management. Integration of Convex and MySQL

ensured fast and structured data storage tailored to each module's needs. The system's AI integrations significantly improved user support and business decision-making, especially for non-tech-savvy retailers.

Challenges included tuning OCR accuracy for low-quality receipts and ensuring consistent user authentication across modules. These were mitigated using preprocessing techniques and centralized authentication via Clerk. Payment gateway integration required handling varied transaction flows, but early integration reduced friction during implementation.

IX.CONCLUSION

RetailEdge demonstrates a scalable, AI-enhanced platform tailored for the unorganised retail sector. By unifying transaction management and receipt digitization, it bridges the digital gap for small retailers. The project's modular design, robust tech stack, and AI integration position it for real-world deployment, with future potential for multilingual support, loyalty systems, and advanced analytics.

X.REFERENCES

- [1] S. Patel, R. Kumar, and A. Mehta, "Adoption of POS Systems by Small Retailers in Developing Economies," *International Journal of Retail Technology*, vol. 8, no. 2, pp. 45–53, 2019.
- [2] A. Gupta and N. Sharma, "Challenges in POS Implementation in Unorganised Markets," *Indian Journal of Digital Economics*, vol. 12, no. 1, pp. 33–40, 2020.
- [3] V. Rao and A. Srinivasan, "A Comparative Study on Open-Source POS Systems for Small Businesses," *Journal of Open Source Software Engineering*, vol. 5, no. 4, pp. 22–29, 2021.
- [4] R. Smith, "An Overview of the Tesseract OCR Engine," in *Proc. International Conference on Document Analysis and Recognition (ICDAR)*, 2007, pp. 629–633.
- [5] H. Li, F. Wang, and Y. Zhang, "Enhancing OCR Accuracy for Financial Documents using Pre-processing Techniques," *IEEE Transactions on Document Analysis and Recognition*, vol. 29, no. 3, pp. 214–222, 2022.
- [6] P. Mehta and R. Bansal, "Context-Aware OCR for Retail Digitization in Emerging Markets," *Journal of Information Systems Research*, vol. 14, no. 1, pp. 17–25, 2022.
- [7] M. N. Nair and D. Joseph, "AI Integration in Retail POS Systems: A Case Study on Predictive Analytics," *International Journal of Artificial Intelligence Applications*, vol. 10, no. 2, pp. 80–88, 2023.
- [8] L. Chen and K. Huang, "Improving Inventory Management using Real-Time Analytics in POS Platforms," *Retail Data Science Journal*, vol. 7, no. 3, pp. 95–103, 2021.
- [9] J. Zhang, A. Lee, and S. Tan, "Serverless Architectures for Scalable Retail Applications," *ACM Cloud Computing Review*, vol. 6, no. 1, pp. 28–36, 2020.
- [10] D. Patel, R. Shah, and Y. Goel, "User Experience Design for Non-Tech-Savvy Retailers: Insights from a Field Study," *Human-Centered Computing Journal*, vol. 13, no. 2, pp. 67–75, 2022.