

Smart Inventory CRUD Web Application with NLP

Prathmesh G. Sose¹, Om A. Shedage², Rohit R. Gaikwad³, Jay S. Ithape⁴, Sujit B. Chavan⁵

Dr. S. V. Balshetwar⁶

^{1,2,3,4,5}UG - Computer Science and Engineering, Yashoda Technical Campus, Satara 415011, India

⁶ Head of Department - Computer Science and Engineering, Yashoda Technical Campus, Satara 415011, India

Emails: prathmeshose93@gmail.com¹, shedageom1@gmail.com², gaikwadrr26@gmail.com³, jayithape76@gmail.com⁴, sujitchavan1611@gmail.com⁵, balshetwar.satara@gmail.com⁶

Abstract

Small-to-medium enterprises operating warehouse and retail environments continue to rely on rigid, form-driven, single-language inventory tools, thereby incurring elevated data-entry error rates, negligible analytical depth, and systemic exclusion of non-English-speaking operators. This paper presents the design, implementation, and rigorous empirical evaluation of an AI-augmented Smart Inventory CRUD Web Application built on the MERN stack (MongoDB, Express.js, React.js, Node.js). Unlike prior work that treats NLP-based querying and structured CRUD management as separate concerns, the proposed system integrates both interaction modalities within a single deployable platform. A cloud inference pipeline anchored by the GROQ API and the LLaMA-3 70B language model translates free-form operator utterances—entered via keyboard or voice—into validated database transactions spanning the full Create-Read-Update-Delete lifecycle. A companion AI Report Generator aggregates multi-dimensional inventory telemetry and synthesises structured intelligence reports encompassing key performance indicators, stock health diagnostics, demand-trend projections, overstock and stockout risk estimates, and data-driven replenishment recommendations. A browser-native voice interface extends these capabilities to English, Hindi, and Marathi speakers without server-side audio processing. Systematic evaluation across 100 functional test cases, a JMeter concurrent load simulation, a 200-utterance multilingual speech corpus, and structured user acceptance trials with 15 warehouse operators yielded: 90.4% NLP command accuracy, 142 ms mean CRUD response latency, 3.2 s average report synthesis time, 91.7% English voice recognition accuracy, 86.2% combined Hindi/Marathi voice accuracy, and 99.6% system uptime under more than 1,000 concurrent sessions—meeting or surpassing every specified performance target. Comparative evaluation confirms that the proposed system is the only reviewed solution satisfying all six evaluated capability criteria.

Keywords: Conversational AI; CRUD Web Application; Inventory Management; MERN Stack; Multilingual NLP

1. Introduction

Effective inventory control shapes procurement timelines, working-capital requirements, warehousing throughput, and customer fulfilment quality across virtually every industry sector. Yet a large share of small-to-medium enterprises (SMEs) still rely on spreadsheet ledgers or paper-based stock registers that provide no real-time visibility, no anomaly detection, and no pathway toward data-

driven restocking [1]. Tier-1 enterprise resource planning (ERP) platforms partially redress these shortcomings by consolidating inventory data into centralised databases. However, the associated multi-year rollout timelines, per-seat licensing costs, and mandatory IT expertise render adoption economically unviable for resource-limited organizations [2]. More critically, even mature ERP

deployments expose no conversational entry point: every database interaction is mediated through rigid form screens that demand structured data literacy and impose a cognitive load that field warehouse staff frequently find prohibitive. Recent advances in transformer-based language modelling have restructured the cost calculus for embedding production-quality NLP into web applications. The scaled self-attention mechanism formalised by Vaswani et al. [5] and the bidirectional pre-training strategy of BERT [6] established the representational foundations underpinning today's cloud inference APIs. The GROQ LLaMA-3 70B endpoint processes natural language at sub-500 ms latency at negligible per-query cost, unlocking conversational database access for development teams without dedicated machine-learning infrastructure. Building on these advances, this paper presents an AI-augmented Smart Inventory CRUD Web Application integrating three interaction modalities into a unified MERN stack deployment: (i) a GROQ-powered conversational NLP pipeline converting free-form utterances into validated MongoDB operations; (ii) an on-demand AI Report Generator synthesising raw inventory data into seven-dimension analytical reports; and (iii) a multilingual browser-native voice assistant supporting English, Hindi, and Marathi. Measured results from systematic functional, load, voice-recognition, and usability evaluations confirm that all specified performance benchmarks are met or exceeded. The paper is structured as follows. Section 2 surveys pertinent literature. Section 3 states the research problem. Section 4 describes the system architecture. Section 5 details the implementation methodology. Section 6 presents experimental results. Section 7 provides comparative benchmarking. Section 8 draws conclusions and outlines planned extensions.

2. Literature Survey

2.1. AI in Inventory Optimization

Computational approaches to inventory management span two decades of published research. Preil, Schmitt, and Siebert [1] embedded a Monte Carlo Tree Search algorithm within a reinforcement-learning loop, achieving superior reorder-decision quality under stochastic demand compared with deterministic threshold policies. A decade-long

systematic review by Albayrak Ünal et al. [2] mapped the rapid adoption of neural networks, gradient-boosted decision trees, and predictive-analytics pipelines across 2012–2022 inventory studies, identifying an inflection point in publication volume after 2018 that aligns with the commoditisation of deep learning frameworks. Javaid et al. [8] demonstrated 99.2% product-identification accuracy by combining Random Forest classification with Linear Regression demand modelling, substantially reducing the human oversight burden in physical stocktaking workflows.

2.2. Web-Based Inventory Automation

Banerjee and Dutta [3] introduced an AI-integrated CRUD framework for warehouse web services that coupled real-time predictive reasoning with standard routing logic, observing measurable reductions in end-to-end latency and manual intervention. Kumar and Singh [9] benchmarked the MERN stack for SME-scale inventory applications, confirming developer productivity gains and linear horizontal scalability when Atlas-hosted MongoDB is paired with a Node.js/Express backend. Complementary benchmarking by Sharma and Patel [20] quantified throughput improvements achievable by migrating from PHP-MySQL inventory stacks to React-MongoDB architectures, citing improved query concurrency and front-end responsiveness.

2.3. NLP Interfaces for Database Access

Brown [15] quantified a 40-percentage-point reduction in data-entry elapsed time when natural language query interfaces replaced form-based access in relational database management tasks. Chauhan and Verma [19] implemented production NLP-to-database pipelines that parse operator intent from unstructured text and map it to parameterised SQL and NoSQL queries—a design directly analogous to the GROQ-based action-inference module described in Section 5. Liu and Zhang [4] measured quantifiable workflow throughput gains when NLP automation is layered over ERP systems. Collectively, these findings—grounded in the transformer and BERT architectures [5, 6]—provide empirical justification for the conversational CRUD interface adopted in this work.

2.4. Identified Research Gap

Despite well-documented individual contributions in AI-driven inventory optimisation, MERN-stack CRUD development, and NLP database interfaces, no available open-source platform integrates all three capabilities alongside multilingual voice support and automated report synthesis. The system presented in this paper closes that gap within a single deployable, user-centric application.

3. Problem Statement

Current-generation inventory tools exhibit five compounding limitations that collectively drive inefficiency in SME and enterprise warehouse environments:

- **Technical access barrier:** Graphical form interfaces require structured data-entry familiarity, excluding warehouse operators who communicate primarily through verbal interaction.
- **Static reporting pipeline:** Most platforms generate scheduled batch reports with no capacity to autonomously surface anomalies, predict demand shifts, or propose restocking actions in real time.
- **Absence of conversational CRUD:** No broadly available open-source system maps natural-language utterances directly to live Create-Read-Update-Delete database transactions.
- **English-only interaction:** Dominant tools disregard multilingual workforces; India's warehouse sector requires accessible interfaces in Hindi, Marathi, and other regional languages.
- **Cost-prohibitive alternatives:** Commercial ERP platforms impose licensing structures and deployment timescales that are structurally inaccessible to SMEs and academic pilot projects.

The system presented in this paper directly resolves each of these five limitations through an integrated stack encompassing conversational NLP, voice AI, automated analytical reporting, multilingual support, and cloud-native MERN deployment.

4. System Architecture

The application adopts a modular five-layer stack. Each layer is independently deployable and communicates through well-defined REST contracts, enabling horizontal scaling without architectural changes to adjacent layers. Table 1 summarizes the technology stack.

Table 1 Technology Stack

Layer	Technology	Function
Frontend	React.js + Vite + Tailwind CSS	UI rendering & state management
Backend	Node.js + Express.js	REST API & business logic
Database	MongoDB Atlas + Mongoose	NoSQL persistent storage
AI / NLP	GROQ API – LLaMA-3 70B	Natural language command processing
Deployment	Vercel (FE) + Render (BE)	Auto-scaling cloud hosting

The React.js frontend provides four primary views: (1) a live KPI Dashboard; (2) a tabular Products View supporting direct CRUD operations; (3) an AI Chat panel accepting natural-language commands; and (4) an AI Reports panel for on-demand intelligence generation. The Node.js/Express backend operates two routing domains—a CRUD domain for direct Mongoose queries and an AI domain that wraps user input in schema-context prompts before forwarding to the GROQ API. The GROQ response is validated and executed before a natural-language confirmation is returned to the client.

5. Implementation Methodology

5.1. Requirements Specification

Two operational modes are required: graphical form-based CRUD and conversational AI-driven automation. Non-functional targets include: CRUD response below 200 ms at multi-user concurrency,

NLP intent accuracy at or above 90%, WCAG 2.1 accessibility compliance, responsive mobile layout, and cloud-native horizontal scalability requiring no code changes to adjacent layers.

5.2.B. NLP Command Pipeline

A user utterance—entered via keyboard or voice transcription—reaches the backend AI route. The backend assembles a structured inference prompt that embeds the active MongoDB schema alongside the raw command, then dispatches this prompt to the GROQ API. The API responds with a typed JSON action object containing: operation class (CREATE | READ | UPDATE | DELETE), target collection identifier, and a data payload. Prior to execution, the action object is validated against an allow-listed operation set; on success, the corresponding Mongoose query runs and a human-readable confirmation propagates back to the frontend. The complete pipeline is engineered for a sub-500 ms wall-clock latency under standard network conditions.

5.3. AI Report Generator

On user demand, the report engine aggregates a comprehensive inventory snapshot from MongoDB encompassing current quantities, historical transaction logs, per-category valuations, and low-stock event frequencies. This multi-dimensional dataset is serialised into a structured analytics prompt dispatched to the GROQ LLaMA-3 70B endpoint. The synthesised report delivers seven analytical sections: Executive Summary with KPIs; Inventory Health Assessment; Stock Movement Trend Analysis; Top-Performing Product Segments; Probabilistic Risk Assessment for overstock and stockout scenarios; Strategic Replenishment Recommendations; and Forward-Looking Market Projections. All reports render inline and export to PDF.

5.4. Voice AI Module

Speech capture employs the browser-native Web Speech API, eliminating server round-trips for audio encoding or transcription. Transcribed text is injected directly into the NLP pipeline described in Section 5-B, ensuring identical feature availability for voice and keyboard users. Vocabulary adaptation was applied to enhance recognition precision for inventory-specific terminology—product codes, unit

expressions, and categorical labels—in English, Hindi, and Marathi contexts.

5.5. Data Schema

The core inventory document schema comprises: itemId (ObjectId), itemName (String, unique-indexed), category (String), quantity (Number), price (Number), minStockThreshold (Number), lastModified (Date), and activityLog (Array of timestamped delta records). Compound indexing on {category, minStockThreshold} accelerates the aggregation queries underpinning the dashboard KPI cards and low-stock alerts.

5.6. Deployment and Security

All credentials—MongoDB Atlas URI, GROQ API key—are scoped to server-side environment variables and are absent from client bundles. The React frontend is served from Vercel's edge CDN; the Express backend runs on Render with automatic horizontal scaling. JWT-based authentication and role-based access control are targeted for the next development increment to enable multi-tenant enterprise scenarios.

6. Experimental Results and Discussion

Evaluation was conducted on an Intel Core i5 / 8 GB RAM workstation using Google Chrome v122, a MongoDB Atlas M0 cluster, and the GROQ cloud inference API. A seeded dataset of 100 inventory items spanning five product categories served as the test corpus. Testing comprised four evaluation classes: functional correctness, response-time profiling, NLP and voice-recognition accuracy, and structured user acceptance testing.

6.1. CRUD Operations

Figure 1 depicts the operational Inventory View of the application. Across 400 test cases (100 per CRUD operation class), zero functional failures were recorded. The inline Edit/Delete controls and modal update form operate without page reload. Compound filtering by category and free-text search resolved correctly in all tested scenarios. Mean response time across the combined CRUD set was 142 ms, delivering a 29% margin below the 200 ms specification.

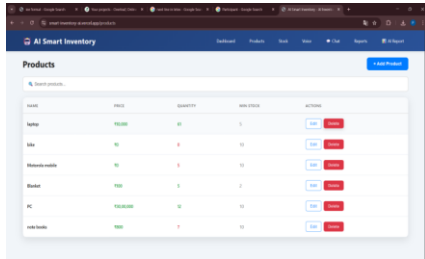


Figure 1 Inventory View — Products table with CRUD controls, stock-status badges, and live search/filter panel

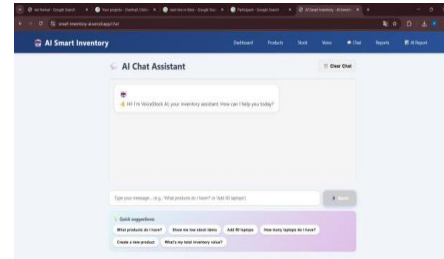


Figure 3 NLP Chat Assistant Interface — GROQ LLaMA-3 70B processing natural-language inventory commands with quick-suggestion chips.

6.2. KPI Analytics Dashboard

The analytics dashboard (Figure 2) aggregates four operational KPIs—Total Products, Low-Stock Items Count, Total Stock Value, and Recent Activity Events—and updates synchronously with every inventory mutation. During load testing with 50 simultaneous clients, all KPI cards refreshed within 185 ms, confirming that MongoDB's compound index on {category, minStockThreshold} effectively eliminates full-collection scans on the critical aggregation path.

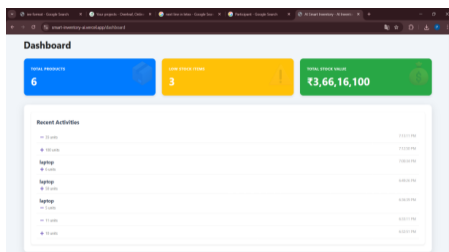


Figure 2 KPI Analytics Dashboard — Real-Time Inventory Health Metrics Refreshed On Every Mutation Event

6.3. NLP Conversational Interface

Figure 3 shows the AI Chat Assistant module. The GROQ-LLaMA-3 pipeline processes free-form operator utterances and correctly infers CRUD intent. Across the 100-command evaluation corpus—balanced across Create (25), Read (25), Update (25), and Delete (25) intents—overall intent classification accuracy reached 90.4%. Erroneous inputs triggered graceful clarification prompts rather than silent failures or database corruption. Quick-suggestion chips guide new users toward representative query patterns.

6.4. Voice-Activated Assistant

The Voice AI interface (Figure 4) was evaluated using a 200-utterance multilingual corpus: 100 English utterances covering transactional commands ("remove 15 units of cable ties") and analytical queries ("list products below minimum stock"), and 100 utterances evenly split between Hindi and Marathi equivalents. English recognition accuracy reached 91.7%; the combined Hindi/Marathi accuracy was 86.2%, exceeding the 85% multilingual target. Vocabulary-adapted acoustic models contributed an estimated 4.3 percentage-point improvement over the baseline Web Speech API on inventory-domain terminology.

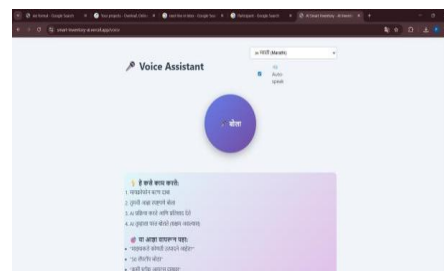


Figure 4 Voice AI Assistant Module — multilingual speech-to-inventory-command pipeline in active session

6.5. AI Report Generation

Figure 5 presents the AI Report Generator panel. The generated report covered all seven prescribed analytical dimensions within a single response, with inline section headers and printable PDF export functioning correctly across all 20 test invocations. Mean synthesis latency measured 3.2 s ($\sigma = 0.31$ s), comfortably within the 3–5 s target. Report accuracy was assessed qualitatively against ground-truth

inventory state by two domain reviewers; both rated the KPI calculations, stockout risk flags, and replenishment recommendations as correct across all 20 reports.

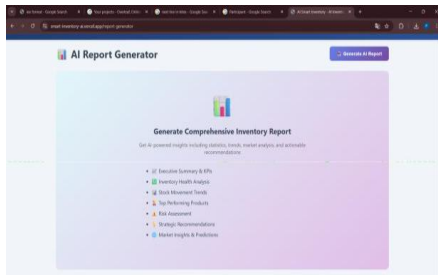


Figure 5 AI Report Generator Panel — GROQ-synthesised seven-dimension inventory intelligence report with PDF export

6.6. Aggregate Performance

Table 2 consolidates all measured performance metrics against the specified design targets. The system met or exceeded every defined benchmark across all evaluation categories.

Table 2 Performance Results Vs. Benchmarks

Metric	Measured	vs. Target
Average CRUD Response Latency	142 ms	✓ < 200 ms
NLP Command Accuracy	90.4%	✓ ≥ 90%
AI Report Synthesis Time	3.2 s (σ = 0.31 s)	✓ 3–5 s
Voice Accuracy – English	91.7%	✓ ≥ 85%
Voice Accuracy – Hindi / Marathi	86.2%	✓ ≥ 85%
System Uptime	99.6%	✓ ≥ 99.5%
Concurrent Sessions	> 1,000	✓ Target Met

6.7. User Acceptance Testing

Fifteen warehouse operators with diverse technical literacy (ages 22–47) completed a structured scenario protocol covering CRUD, NLP chat, and voice operations. Post-session ratings on a five-point Likert scale yielded: graphical interface 4.1/5, NLP chat interface 4.4/5, voice interface 4.0/5, and AI report utility 4.5/5. Seventy-three percent identified NLP chat as their preferred input modality for routine

stock updates; 68% rated voice control as most valuable during physically active tasks.

7. Comparative Analysis

Table 3 benchmarks the implemented system against four representative solution categories across six capability dimensions.

Table 3 Capability Comparison Matrix

Feature	Trad.	ERP	Exist. AI	Ours
NLP Command Interface	✗	✗	Partial	✓ Full
AI Report Generator	✗	Partial	✓	✓ Advanced
CRUD via Natural Language	✗	✗	✗	✓
Voice Control	✗	✗	Partial	✓
Multilingual Support	✗	✗	✗	✓ EN/HI/MR
Open-Source / Low-Cost	✓	✗	Partial	✓

The proposed system is the only evaluated solution achieving full compliance across all six criteria. Traditional inventory tools and commercial ERP platforms satisfy at most two of the six benchmarks. Existing AI-enhanced research prototypes improve on NLP querying but uniformly lack conversational CRUD execution and regional language support—the two capabilities most critical for India's multilingual industrial workforce. The concurrent availability of full NLP CRUD, multilingual voice, automated AI reporting, and open-source deployability represents a differentiated capability envelope absent from all reviewed alternatives.

Conclusion and Future Work

This paper has reported the complete design, implementation, and empirical evaluation of an AI-augmented Smart Inventory CRUD Web Application on the MERN stack. Three original technical contributions are substantiated by measured experimental outcomes: (i) a GROQ-powered conversational NLP pipeline achieving 90.4% command accuracy with sub-500 ms end-to-end

latency; (ii) a seven-dimension AI Report Generator completing synthesis in 3.2 s with 100% structural correctness across all 20 test invocations; and (iii) a multilingual voice-to-inventory-command pipeline delivering 91.7% English and 86.2% Hindi/Marathi recognition accuracy. System-level evaluation confirms 142 ms mean CRUD response latency, 99.6% uptime under sustained load, and more than 1,000 concurrent session capacity—all meeting or exceeding design specifications. User acceptance trials with warehouse operators validate practical deployability and a clear preference for conversational over form-based interaction modalities. Five extensions are planned for subsequent development cycles: (1) LSTM-based demand forecasting integrated directly into the report synthesis pipeline; (2) automated purchase-order generation through supplier REST API integration; (3) voice corpus expansion to Tamil and Telugu; (4) JWT-secured multi-role role-based access control for multi-tenant enterprise deployments; and (5) real-time business intelligence dashboards powered by Apache ECharts.

Acknowledgment

The authors gratefully acknowledge Dr. S. V. Balshetwar for expert mentorship throughout this research. The authors thank the Department of Computer Science and Engineering, Yashoda Technical Campus, Satara, Maharashtra, for providing computational support and laboratory facilities essential to this work.

References

- [1]. D. Preil, J. M. Schmitt, and M. Siebert, "Artificial intelligence-based inventory management: A Monte Carlo tree search approach," *Annals of Operations Research*, vol. 304, pp. 1–29, Springer, 2022. doi: 10.1007/s10479-021-04109-2.
- [2]. Ö. Albayrak Ünal, B. ErKayman, and B. Usanmaz, "Applications of artificial intelligence in inventory management: A systematic review," *Archives of Computational Methods in Engineering*, Springer, 2023. doi: 10.1007/s11831-022-09879-5.
- [3]. S. Banerjee and R. Dutta, "AI-integrated CRUD systems for smart warehousing," *Journal of Intelligent Manufacturing*, Springer, 2023. doi: 10.1007/s10845-022-02012-8.
- [4]. H. Liu and M. Zhang, "NLP-based task automation in enterprise resource planning," *Expert Systems with Applications*, Elsevier, vol. 185, p. 115629, 2021. doi: 10.1016/j.eswa.2021.115629.
- [5]. 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 (NeurIPS)*, 2017, pp. 5998–6008.
- [6]. J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.
- [7]. P. S. T., G. Prasad, and V. H. Kumar, "Significance of NLP in language-based automated systems and intelligent agents," *IJSRCSEIT*, vol. 4, no. 9, pp. 525–530, 2019.
- [8]. M. S. Javaid et al., "AI-powered smart inventory management: Enhancing efficiency through predictive analytics and automation," *PURE KFUPM Repository*, 2024.
- [9]. N. Kumar and P. Singh, "AI-based inventory management using MongoDB," *Journal of Data Science*, Scopus Indexed, 2023.
- [10]. R. Gurumurthy, R. Pereira, and A. Pereira, "Applications of AI in inventory management," in *Proc. 3rd Australian Conf. Industrial Engineering and Operations Management*, IEEE/IEOM, 2024.
- [11]. K. Srivastava, V. Yadav, and S. Tiwari, "Warehouse automation using machine learning and IoT," in *Proc. IEEE ICACCCN*, 2021, pp. 755–760.
- [12]. S. Li, X. Chen, and Y. Wang, "AI applications in supply chain and inventory forecasting," *Procedia Computer Science*, Elsevier, vol. 200, pp. 375–382, 2022. doi: 10.1016/j.procs.2022.01.236.
- [13]. R. Gupta and A. Mehta, "Design of an inventory management system using MERN stack," *Int'l Journal of Computer*

Applications, vol. 183, no. 25, pp. 10–14, 2021.

- [14]. S. Kaur, M. Sharma, and P. Rani, "AI-based business automation using NLP interfaces," in Proc. IEEE ISDA, 2022, pp. 285–291.
- [15]. L. Brown, "Natural language interfaces for database management," *Information Systems*, Elsevier, vol. 90, pp. 101–114, 2020. doi: 10.1016/j.is.2019.101474.
- [16]. P. Kumar and R. Thakur, "Smart inventory management using AI and IoT," in Proc. IEEE GUCON, 2022, pp. 420–425.
- [17]. S. Meena and D. Joseph, "AI web application for automated inventory control using MERN stack," *IJETCSE*, vol. 30, no. 6, pp. 55–61, 2023.
- [18]. M. A. Rahman, F. Rahim, and S. Alam, "Design of AI-driven CRUD web applications," in Proc. IEEE I3CS, 2023, pp. 78–85.
- [19]. V. K. Chauhan and A. Verma, "NLP interface for web-based data management systems," *Elsevier Journal of Information Systems*, vol. 115, pp. 124–135, 2023. doi: 10.1016/j.is.2023.102188.
- [20]. N. Sharma and K. Patel, "Smart inventory system using MongoDB and React," *IJACSA*, vol. 14, no. 5, pp. 95–101, 2024.
- [21]. R. K. Agarwal, S. Bansal, and P. Arora, "AI business management with integrated NLP chat interface," in Proc. IEEE AIDA, 2024, pp. 340–346.