Jurnal Pendidikan Teknologi Kejuruan, Vol. 9, No. 1, pp. 1-20, February 2026
e-ISSN 2621-1548; p-ISSN 2621-3273

Implementation of a WhatsApp-based Canteen Menu Chatbot using TF–IDF and cosine similarity: A case study of employees at Matahari Department Store, Sunrise Mall and Mojokerto

Ramadhani Handika Saputra* **and Hadi Sucipto**

Department of Informatics Engineering, Faculty of Information Technology, Universitas Hasyim Asy'ari, **Indonesia**

*Corresponding author: ramadhanisaputra@mhs.unhasy.ac.id
Received: 11 November 2025; *Revised:* 12 January 2026; *Accepted:* 15 January 2026

Cite this https://doi.org/10.24036/jptk.v9i1.48523

Abstract: This study addresses the limitations of traditional workplace canteen menu information sharing, such as noticeboards and verbal announcements, which can cause delays in shift-based settings. A WhatsApp-based canteen menu chatbot was designed, implemented, and evaluated for Matahari Department Store Sunrise Mall Mojokerto (Indonesia) using Cosine Similarity for text matching. The system integrates WhatsApp messaging via the WhatsApp API with an application server and an admin web panel built on Laravel–MySQL, enabling the management of daily menus, canteen schedules, FAQ patterns, user access, and conversation logs. User queries and stored question patterns were preprocessed (lowercasing, cleaning, tokenization), vectorized using TF–IDF, and matched using Cosine Similarity with a 0.6 acceptance threshold to route responses into three classes (Static, Menu, or Schedule) or a fallback reply. Functional verification through black-box testing confirmed the correct operation of core modules and end-to-end message handling. Similarity computation was validated against a manual example (0.816 vs. 0.8165). Robustness testing on 50 employee questions (25 standard and 25 informal/misspelled) achieved 92% accuracy (46/50); retrieval effectiveness measured as micro-F1 across the three response classes was approximately 0.95, with four queries treated as unmatched. A baseline comparison showed that Cosine Similarity was more tolerant to paraphrasing than SQL keyword matching. The case study ran from August to November 2025, and conversation logs were used during the maintenance phase to refine FAQ patterns and improve coverage. Overall, the chatbot offers a lightweight and deployable solution for routine, time-sensitive internal information services in retail workplaces.

Keywords: WhatsApp Chatbot; TF–IDF; cosine similarity; information retrieval; workplace information service; canteen menu

1. Introduction

The rapid advancement of information technology has encouraged organizations to provide information services that are responsive, accurate, and easily accessible (He et al., 2021; Sofyani et al., 2020; Wang & Wu, 2021). One increasingly adopted approach is the use of chatbots as conversational agents that automatically deliver information and answer user queries, enabling institutions to improve service efficiency while reducing reliance on human operators (Andrade & Tumelero, 2022; Mohamad Suhaili et al., 2021). Chatbots have been implemented across education, healthcare, finance, and public services due to their ability to streamline communication and support users in real time (Branda et al., 2025; Chen et al., 2024; Laymouna et al., 2024).

However, many workplace settings still face a practical operational issue: canteen menu information is frequently disseminated manually via noticeboards or verbal announcements (Breathnach et al., 2020; Taylor et al., 2021). Such methods are vulnerable to delayed updates, limited reach, and inconsistent information, especially in shift-based workplaces where employees may not be present when announcements are posted. This condition is also observed in the employee canteen of Matahari Department Store Sunrise Mall Mojokerto, where menu information dissemination has not been optimized through digital services, contributing to inefficiencies and delayed access to daily menu updates (Hayati, 2019). In fast-paced service environments, improving internal information delivery is not a minor administrative task; it directly shapes employee convenience and everyday operational routines (Alateeg & Alhammadi, 2024; Aristovnik et al., 2024).

Recent studies indicate that WhatsApp is a particularly strategic channel for organizational information services in Indonesia because it is widely used and does not require users to install additional applications (Himpong et al., 2023; P & G, 2024). Therefore, integrating a chatbot into WhatsApp provides a lightweight, accessible solution for delivering canteen menu information to employees. For accurate responses, the chatbot requires a text-matching mechanism that can identify the closest intent or question pattern from a predefined dataset.

One widely used method in information retrieval and chatbot question matching is Cosine Similarity, which computes similarity between a user query vector and dataset vectors to retrieve the most relevant response (Lokman et al., 2020; Shrivastava et al., 2023). Prior work suggests that Cosine Similarity, often paired with TF–IDF weighting, supports fast and accurate responses in chatbot services. For example, (L. Zhang, 2025) reported effective performance for student services when Cosine Similarity was integrated with TF–IDF and Naïve Bayes, while (Sachdeva et al., 2024) showed that chatbots can efficiently handle repetitive questions and provide critical information to users.

Nevertheless, the literature still shows limited evidence on applying and evaluating Cosine Similarity–based chatbots specifically for workplace canteen menu information in shift-based retail environments. Many studies emphasize academic services or general customer support, whereas fewer examine internal organizational services characterized by routine, time-sensitive content (daily menus) and diverse employee schedules (Adu-Gyamfi et al., 2021; Blagoev et al., 2024; Cedergren & Hassel, 2024; Paul et al., 2024). Moreover, existing research rarely presents and evaluates a complete end-to-end implementation on WhatsApp, from dataset preparation and similarity computation to deployment and performance evaluation, in a real-world workplace context (Amrit & Narayanappa, 2025; Dwivedi et al., 2021; Moganadas & Goh, 2022).

To address this gap, this study proposes developing a WhatsApp-based chatbot for canteen menu information at Matahari Department Store Sunrise Mall Mojokerto, using Cosine Similarity as the core text-matching method. The novelty of this research lies in: (1) implementing Cosine Similarity for menu-information retrieval in an employee canteen context; (2) deploying the service through WhatsApp to maximize accessibility without additional installations; and (3) providing an end-to-end system workflow for handling routine, repetitive, and time-sensitive internal queries in a shift-work environment, including an empirical evaluation of response relevance and access efficiency.

Jurnal Pendidikan Teknologi Kejuruan, Vol. 9, No. 1, pp. 1-20, February 2026
e-ISSN 2621-1548; p-ISSN 2621-3273

Figure 1. Conceptual Framework of WhatsApp-Based Canteen Menu Chatbot Using Cosine Similarity

Figure 1 summarizes the proposed end-to-end conceptual framework for developing and deploying a WhatsApp-based chatbot to deliver canteen menu information in a shift-based workplace environment efficiently. It begins with the identification of the operational problem, manual and inefficient menu dissemination, followed by the technological solution integrating WhatsApp as the communication platform and Cosine Similarity with TF–IDF as the core text-matching mechanism. The framework visualizes the system workflow from dataset preparation, vectorization, and similarity computation to response retrieval, and from deployment to empirical evaluation of response relevance and access efficiency. The model highlights how user queries are processed through similarity scoring to generate the most relevant predefined menu response, ultimately leading to improved accessibility, faster information retrieval, and enhanced employee convenience. Accordingly, this study is guided by the following research questions:

RQ1. How can a WhatsApp-based chatbot be designed and implemented to deliver employee canteen menu information effectively in a shift-based workplace setting?
RQ2. How can Cosine Similarity be applied to match employee questions with a predefined dataset to produce relevant menu information responses?
RQ3. How effective is the proposed chatbot in terms of response relevance and user access efficiency for retrieving canteen menu information?

This study contributes in three ways. In practice, it delivers a deployable WhatsApp chatbot solution that improves the speed and reach of canteen menu information distribution to employees. Methodologically, it provides a replicable workflow for building a Cosine Similarity–based text-matching chatbot (including dataset structuring and similarity-based retrieval) for routine internal information services. Empirically, it adds evidence from a real-world retail workplace case, extending the discussion of chatbot adoption beyond dominant domains such as education and customer service.

2. Methods

2.1. Study design and timeline

This study employed an applied research design, using a software engineering approach, to address the delayed and uneven distribution of daily canteen menu information in a shift-based workplace.

A case study was conducted at the employee canteen of Matahari Department Store Sunrise Mall, Mojokerto, Indonesia, from August to November 2025.

2.2. System development procedure

The system was developed using the Waterfall model (Petersen et al., 2009; Saravanos & Curinga, 2023), as illustrated in Figure 2, which organizes development into five sequential phases: Analysis, Design, Implementation, Testing, and Maintenance. Analysis defined the system requirements through problem identification, a literature review, and field data collection (observations and interviews). Design converted these requirements into the system architecture, database schema, and interaction flow for the WhatsApp chatbot and admin web panel. Implementation built the working system, including the admin panel, API services, and database integration. Testing verified that all functions operated correctly and that the chatbot returned appropriate responses. Maintenance ensured continued reliability after deployment by monitoring performance and updating content, including menu/schedule data and FAQ patterns. In this study, conversation logs collected during deployment were reviewed to identify recurring unmatched variations and to update/add FAQ patterns accordingly, which is treated as content maintenance rather than redevelopment of the core system.

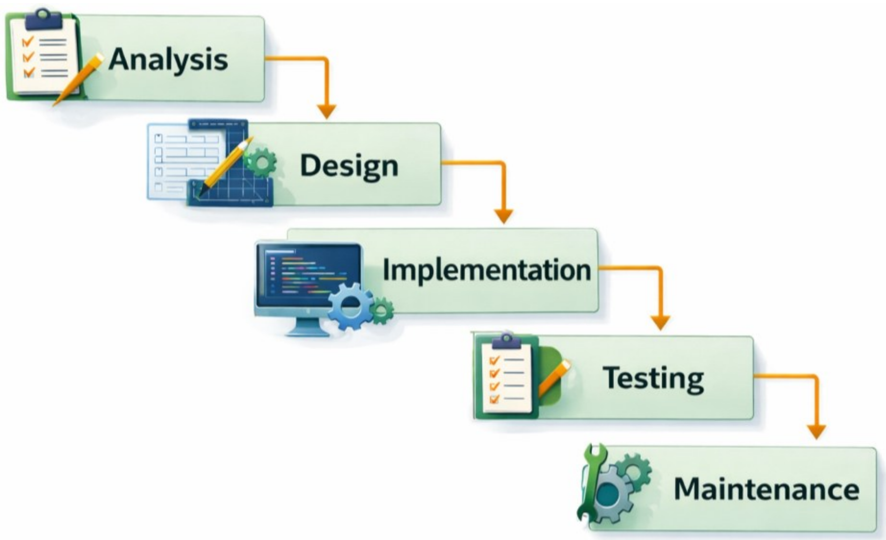

Figure 2. Waterfall model

2.3. Data collection

To define user needs and populate the chatbot knowledge base (Kernan Freire et al., 2025; Skuridin & Wynn, 2024), two techniques were used: (1) direct observation of menu-update routines and staff–employee information exchanges, and (2) semi-structured interviews with canteen staff, employees, and supervisors to identify frequently asked questions, preferred response formats, and communication constraints in a shift-based work context.

2.4. Knowledge base construction

The knowledge base was structured into three content groups: (a) daily menu information, (b) canteen schedule (open/closed), and (c) FAQ patterns representing recurring employee queries. Each FAQ entry stored a question pattern and a response type (menu, schedule, or static) to support both dynamic and fixed responses in Table 1.

Table 1. Knowledge-based structure and response logic for the WhatsApp canteen chatbot

Knowledge base component	Purpose	Minimum fields stored	Response type	Example user query	Example system response
Daily menu	Deliver up-to-date daily menu information	date, menu_text *(optional: menu_items as list/JSON)*	Dynamic (menu)	*"Menu kantin hari ini apa?"*	*"Menu hari ini: nasi, ayam goreng, sayur sop, tempe, teh."*
Canteen schedule	Inform employees whether the canteen is open/closed	date, status *(open/closed) (optional: open_time, close_time)*	Dynamic (schedule)	*"Besok kantin buka tidak?"*	*"Kantin buka besok (07.00–15.00)."*
FAQ – static information	Answer recurring questions that do not depend on the date	question_pattern, answer_text *(optional: keywords)*	Static	*"Bisa bayar pakai apa?"*	*"Metode pembayaran: tunai dan QRIS."*
FAQ – routing rules *(optional but recommended)*	Route queries to the appropriate module (menu/schedule/static)	question_pattern, response_type *(menu/schedule/static) (optional: target_id)*	Router (intent routing)	*"Jam operasional kantin jam berapa?"*	*(routed to schedule module) "Kantin beroperasi pukul 07.00–15.00."*

2.5. System architecture and implementation

The system connected WhatsApp messaging to an application server via a WhatsApp gateway/API. Employee messages were forwarded to the server and matched against the stored knowledge base to generate an automatic reply, which was returned through WhatsApp. A web-based admin panel was implemented to manage menus, schedules, FAQ entries, users, and chat logs, ensuring the knowledge base remained current and auditable. The system was implemented using Laravel (admin interface and API services), MySQL (database), and a WhatsApp gateway service (Figure 3).

Figure 3. WhatsApp-based chatbot system flow

To clarify system roles, Figure 4 presents the use case diagram. Employees interact through WhatsApp to request menus, schedules, and static FAQs, while the canteen admin maintains operational content and reviews logs through the web panel.

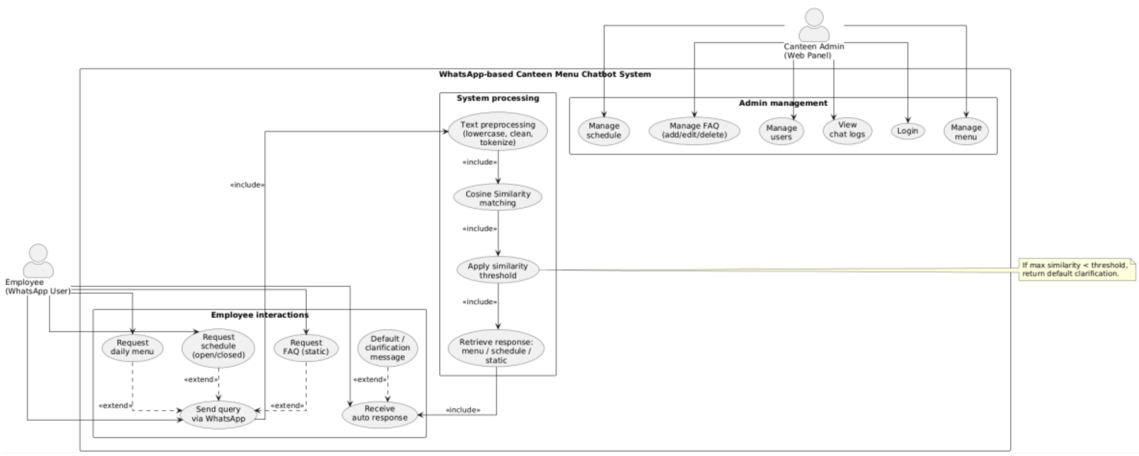

Figure 4. Use case diagram (Admin vs Employee)

2.6. Text preprocessing and cosine similarity matching

User queries and stored question patterns were first standardized through text preprocessing (lowercasing, basic cleaning (e.g., removing punctuation/non-informative symbols), and tokenization). Each processed text was then transformed into a vector using a shared vocabulary with TF–IDF weighting, ensuring consistent representation across employee queries and stored patterns (Liu et al., 2022; W. Zhang et al., 2011). Let q denote the TF–IDF vector of the user query and d denote the TF–IDF vector of a stored question pattern. The similarity between q and d was computed using Cosine Similarity, as shown in Equation (1).

$$\cos(\theta) = \frac{\sum_{i=1}^{n} q_i d_i}{\sqrt{\sum_{i=1}^{n} q_i^2}\sqrt{\sum_{i=1}^{n} d_i^2}} \quad (1)$$

The chatbot returned the response linked to the highest similarity score. To prevent incorrect matches, a similarity threshold was applied; if the best score fell below it, the system treated the query as unmatched and returned a default clarification message.

2.7. Evaluation and performance metrics

Evaluation combined functional verification and retrieval-quality testing (Table 2). Black-box testing validated core functions, including admin management modules and end-to-end chatbot response delivery (Nidhra, 2012). Matching validity was verified by comparing system-similarity outputs with manual Cosine Similarity calculations for selected sample pairs. Robustness was examined through threshold testing using varied employee phrasing to control false matches. Retrieval effectiveness was assessed using Precision, Recall, and F1-score on a labeled set of employee queries. Queries were labeled into three intended response classes (Static, Menu, Schedule). If the system produced the correct class and response type, it was counted as a correct retrieval; otherwise, it was counted as an error. Queries that fell below the 0.6 threshold were treated as "unmatched" and counted as incorrect for the intended class. Micro-averaged Precision, Recall, and F1 were then computed across the three classes.

Table 2. Evaluation protocol and metrics

Evaluation	Purpose	Output metric
Black-box	functional correctness	pass/fail per feature
Manual vs system similarity	calculation consistency	error rate/match consistency
Threshold tolerance	robustness to paraphrase/noise	acceptance rate, false match rate
Retrieval effectiveness	relevance of answers	precision, recall, F1-score
Baseline comparison	added value vs SQL matching	ΔF1 and response-time difference

2.8. Ethical considerations and data privacy

The study used operational information (menus, schedules, and FAQs) and did not require sensitive personal data. Chat logs were used only for debugging and evaluation, with data minimization applied. Access to the admin dashboard and database was restricted to authorized personnel, and any exported evaluation data was anonymized and stored securely for research reporting only.

3. Results

3.1. End-to-end chatbot workflow (cosine similarity)

The WhatsApp-based canteen chatbot functions as an end-to-end retrieval pipeline (Figure 5). The Wablas gateway forwards employee queries submitted through WhatsApp to the application server, where the request is validated, and the message text is processed. The server retrieves candidate entries from the combined knowledge base (FAQ, menu, and schedule), computes Cosine Similarity to identify the best match, and generates a response according to the matched entry type. If the maximum similarity score does not meet the acceptance threshold, the system returns a fallback clarification message rather than producing a low-confidence answer. Each interaction is stored in the chat log, and the final response is delivered to the employee via WhatsApp.

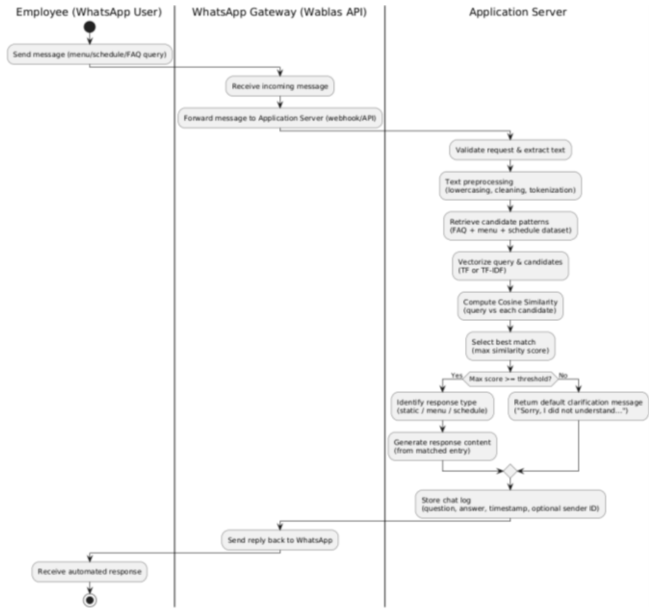

Figure 5. Chatbot workflow using Cosine Similarity (end-to-end flow from WhatsApp query to automated response)

3.2. Response selection and handling unmatched queries

To operationalize routing, the chatbot maps matched queries into three response types: Static, Menu, and Schedule, as illustrated in Table 3 and the decision flow in Figure 6. This structured routing prevents out-of-domain requests, such as product pricing, from being answered incorrectly and ensures that menu/schedule replies are generated from the most current records.

Table 3. Response type selection based on employee questions

Employee question (example)	Response type	System response (example)
"I want to ask, what is the favorite menu in the Matahari canteen?"	Static	"The favorite menu is ayam cemet and ayam kremes."
"What is today's canteen menu?"	Menu	"Today's canteen menu is …"
"Is the canteen open today?"	Schedule	"The canteen is open today."
"How much is a batik shirt?"	Not matched	"Sorry, I do not understand your question."

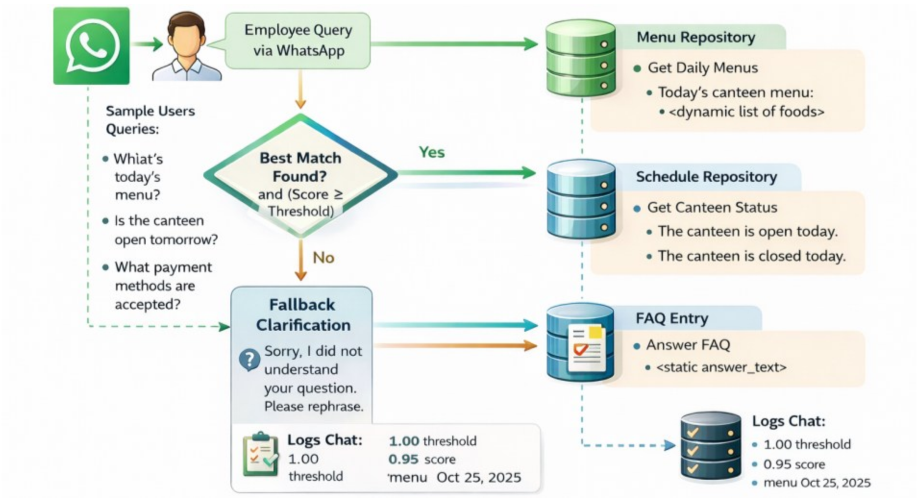

Figure 6. Decision process for selecting the correct response type (static/menu/schedule) or fallback reply

3.3. Conversation logging for monitoring and continuous improvement

For monitoring and continuous improvement, the system records each WhatsApp interaction after responding (Figure 7). The log captures a hashed sender identifier, the user question, the system answer, and a timestamp, with optional matching metadata (e.g., similarity score, threshold, matched/unmatched label). The admin reviews matched and unmatched records in the monitoring interface to identify recurring question variants not represented in the dataset, then updates FAQ patterns accordingly. This closes the improvement loop by grounding dataset expansion in observed usage rather than assumptions.

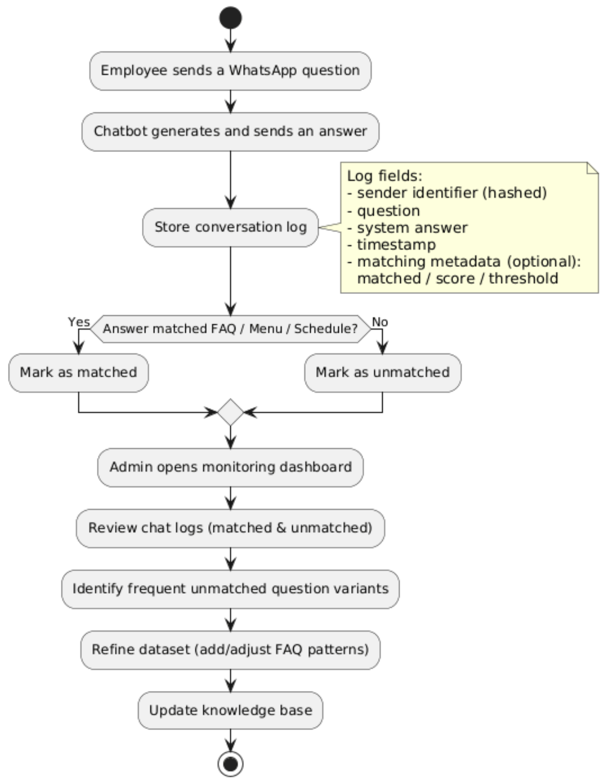

Figure 7. Conversation logging workflow for monitoring and continuous improvement

3.4. System implementation (admin web panel and WhatsApp interface)

The system was implemented as a two-interface solution: a web-based admin panel for maintaining operational content and a WhatsApp chatbot interface for employee access. The admin panel supports content governance by enabling the canteen administrator to update the knowledge base (menus, schedules, and FAQ patterns), manage authorized accounts, and review interaction logs for monitoring and iterative refinement. Employees access the chatbot directly through WhatsApp to request daily menus, check the canteen open/closed status, and ask recurring canteen-related questions in real time. This separation of roles ensures that responses remain current (admin-maintained) while access remains immediate (employee-facing).

Admin web panel (Figures 8–13). Figure 8 presents the dashboard as the main navigation hub, linking to all management modules required to operate the chatbot. Figure 9 shows the chat-log monitoring page, which provides both outcome summaries (e.g., answered vs. failed queries) and message-level records (sender, question, answer, timestamp), allowing the administrator to identify repeated issues and prioritize dataset updates. Figure 10 shows the FAQ management module, where the admin adds and edits question patterns and assigns a response type (Static, Menu, or Schedule), ensuring the matching engine routes queries to the appropriate response logic. Figures 11 and 12 show the schedule and daily menu modules, which store date-specific operational data so the chatbot can return accurate answers that depend on the requested day. Figure 13 shows the user management module, which controls access privileges and supports accountability by limiting dataset and content modification to authorized accounts.

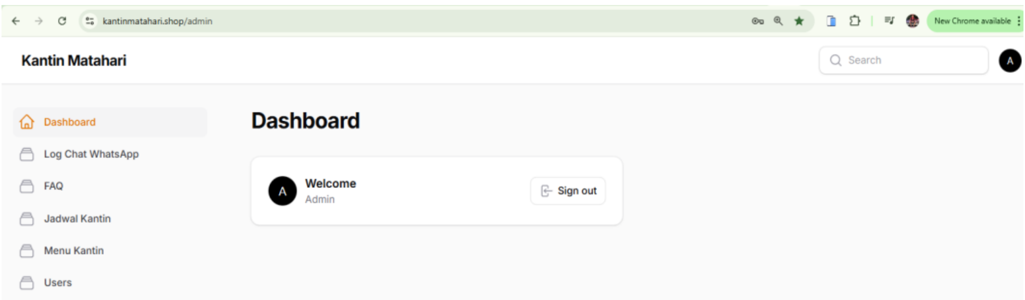

Figure 8. Admin dashboard interface (navigation to core management modules)

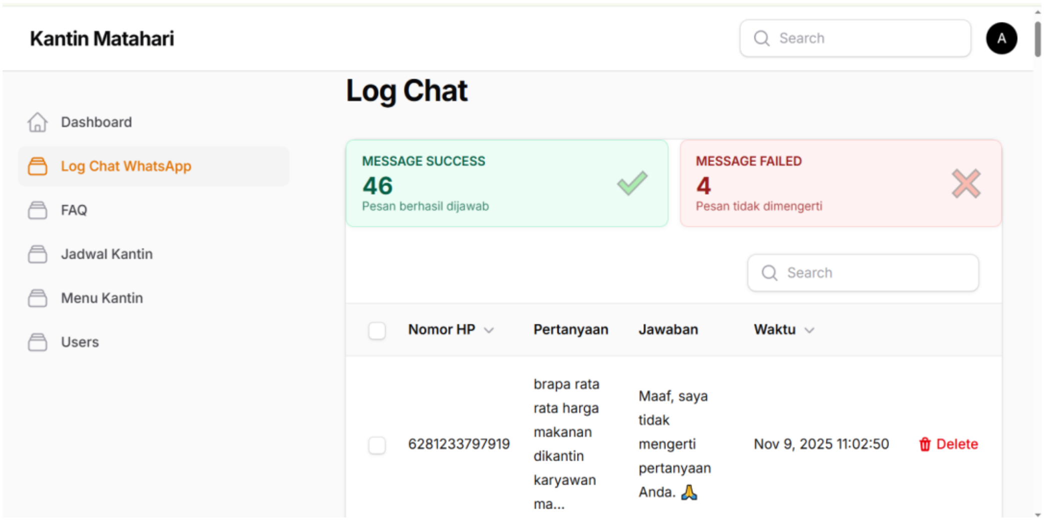

Figure 9. WhatsApp chat log page (conversation history monitoring)

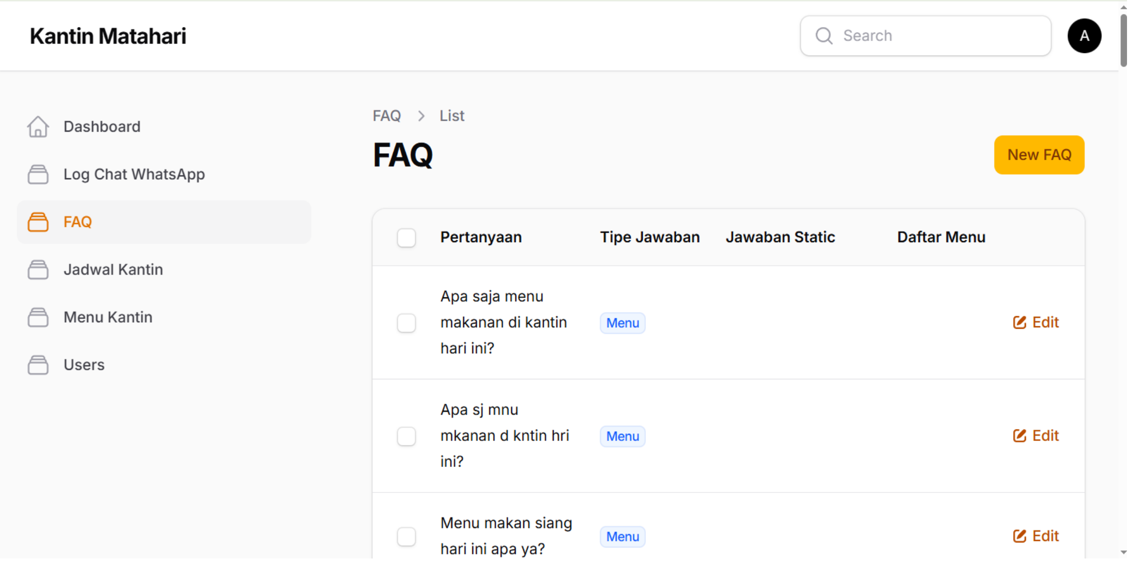

Figure 10. FAQ management page (Frequently Asked Questions)

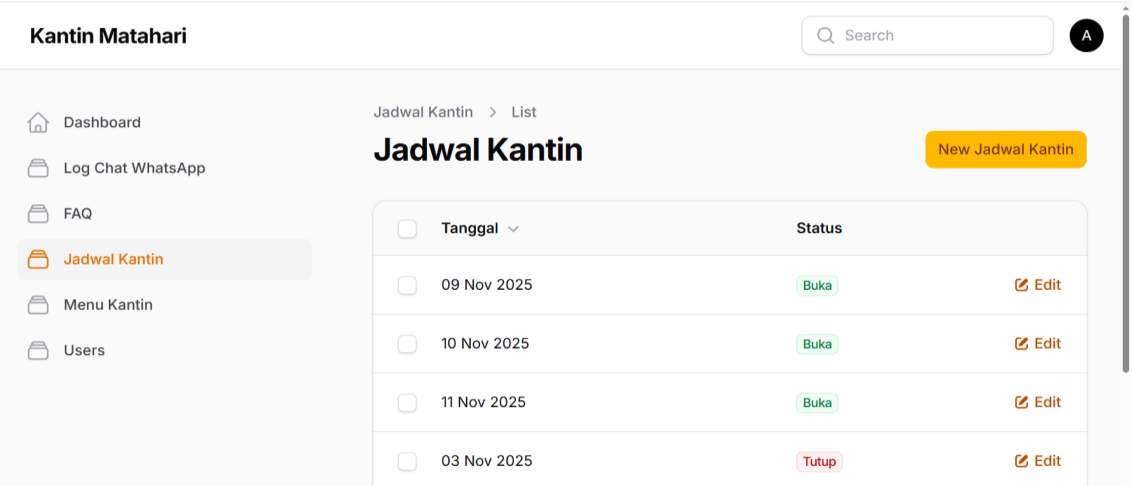

Figure 11. Canteen schedule management page

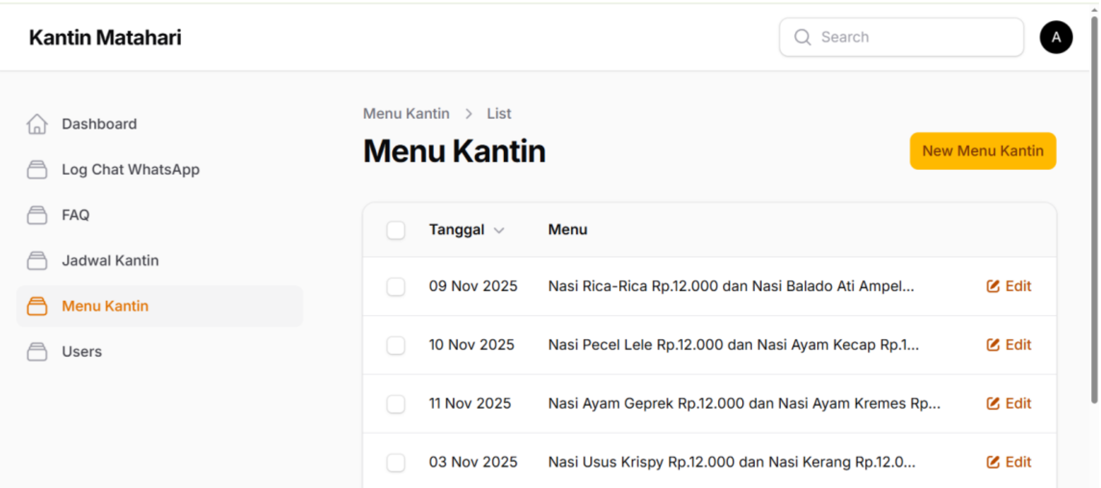

Figure 12. Daily menu management page

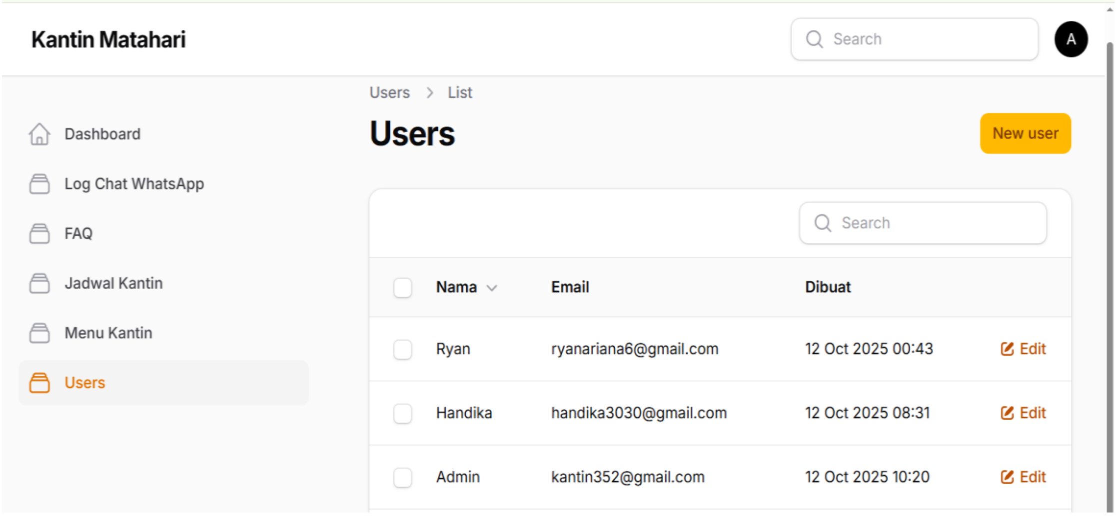

Figure 13. User management page

Figure 14. Example employee–chatbot interaction via WhatsApp (testing view)

Figure 14 demonstrates a multi-turn employee interaction in WhatsApp, where the chatbot responds instantly to sequential canteen-related questions (e.g., today's menu, canteen operating hours, and popular menu items). The exchange illustrates that employees can retrieve routine information through a single familiar channel without requiring additional applications, while the system maintains consistency by returning standardized, task-focused replies.

3.5. Functional verification (black-box testing)

Black-box testing confirmed that the system's main modules produced the expected outputs given valid and invalid inputs. All tested features succeeded, including admin authentication, CRUD operations for menus/schedules/FAQ/users, chat-log recording, and chatbot response handling for matched and unmatched queries, as demonstrated in Table 4.

Table 4. Black-box testing summary

No	Test scenario	Input	Expected output	Result
Admin Login				
1	Admin logs in using a valid email and password	Enter a correct email and password	Redirected to the chatbot website/admin page	Successful
2	Admin logs in using an invalid email and password	Enter an incorrect email and password	"These credentials do not match our records."	Successful
WhatsApp chat log				
1	Employee sends a message to the chatbot	*"Menu kantin hari ini apa?"*	The system stores the sender number, question, and answer in the chat log table	Successful
2	Canteen admin opens the Chat Log page	Click Log Chat menu	System displays the full conversation history	Successful
Frequently asked questions				
1	Admin adds a new question to the FAQ	Click Add FAQ → fill in data	Data is saved to the database and appears in the FAQ list	Successful
2	Admin edits an FAQ entry	Click Edit → modify question/answer	FAQ data updated successfully	Successful
3	Admin deletes an FAQ entry	Click Delete on one FAQ item	FAQ data removed from the database	Successful
Canteen schedule				
1	Admin adds a new schedule entry	Enter the date and set the status to Open	Schedule saved and displayed in the schedule table	Successful
2	Admin updates schedule status	Change status to Closed	Schedule updated and saved	Successful
Daily nenu				
1	Admin adds a new menu entry	Enter today's date and menu data	The menu was saved and displayed in the menu list	Successful
2	Admin edits a menu entry	Click Edit → update menu list	Menu data updated successfully	Successful
3	Admin deletes a menu entry	Click Delete on a menu item	Menu data removed from the database	Successful
Users				
1	Admin adds a new user	Enter name, email, and password	User data is saved and displayed in the user list	Successful
2	Admin edits a user	Click Edit → modify user data	User data updated successfully	Successful
3	Admin deletes a user	Click Delete on a user	User removed from the system	Successful
Chatbot question submission				
1	User submits a question that exists in the database	*"Kantin buka jam berapa?"*	"The canteen opens at 12:00 WIB."	Successful
2	User submits a question that does not exist in the database	*"Baju batik harganya berapa?"*	"Sorry, I do not understand your question."	Successful

3.6. Validation of cosine similarity computation (manual vs. system)

To verify the correctness of the similarity implementation, one sample sentence pair was calculated manually and then compared with the system's output. The test query *"Berapa harga nasi goreng?"* was compared to the dataset pattern *"Harga satu porsi nasi goreng berapa?"*. After preprocessing and term-frequency vectorization, the manually computed Cosine Similarity score was 0.816 (81.6%), whereas the system returned 0.8165 (81.65%), as shown in Table 5 and Figure 15. The near-identical results confirm that the chatbot's similarity computation aligns with the expected mathematical calculation.

Table 5. Manual computation summary

Step	Output (summary)
Preprocessing	lowercase + tokenization
Vocabulary size	6 unique terms
TF vectors	Query: [1,1,1,1,0,0]; Dataset: [1,1,1,1,1,1]
Cosine Similarity	0.816 (manual) vs 0.8165 (system)

```
# ---------- 1. Dataset pertanyaan dan jawaban ----------
dataset = {

    "Harga satu porsi nasi goreng berapa?": "Harga nasi goreng seprosi Rp.12.000"
}

# ---------- 2. Fungsi preprocessing ----------
def preprocess(text):
    # Case folding
    text = text.lower()
    # Hilangkan karakter non alfanumerik
    clean_text = ''.join(ch if ch.isalnum() or ch.isspace() else ' ' for ch in text)
    # Tokenizing
    tokens = clean_text.split()
    return tokens

# ---------- 3. Fungsi cosine similarity ----------
def cosine_similarity(tokens1, tokens2):
    # Gabungkan semua token unik
    all_tokens = set(tokens1 + tokens2)
    vec1 = []
    vec2 = []
```

```
Tanya (atau ketik 'keluar' untuk berhenti): Berapa harga nasi goreng?
Jawaban: Harga nasi goreng seprosi Rp.12.000
Kemiripan: 81.65%

Tanya (atau ketik 'keluar' untuk berhenti):
```

Figure 15. System similarity output (Google Colab simulation)

3.7. Threshold tolerance and robustness

Threshold tolerance testing evaluated the chatbot's robustness to variations in phrasing and spelling. A similarity threshold of 0.6 was applied: when the highest similarity score fell below this value, the system returned a fallback response ("Sorry, I do not understand your question"). Out of 50 test queries, the chatbot answered 46 correctly and failed to match 4, indicating strong overall robustness with weaknesses mainly in informal or misspelled inputs. The response distribution by input type is shown in Table 6.

Table 6. Response success by input type (n = 50)

Question type	Total	Correct answers	Incorrect/unmatched
Correct writing	25	25	0
Informal/misspelled writing	25	21	4
Total	50	46	4

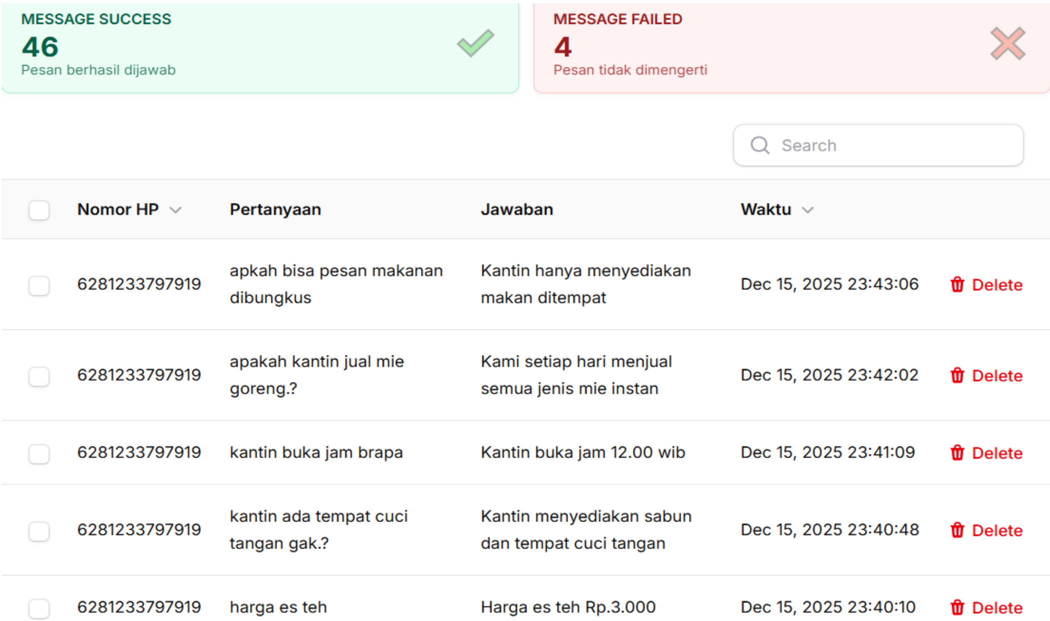

Figure 16. Message Success view (46 answered)

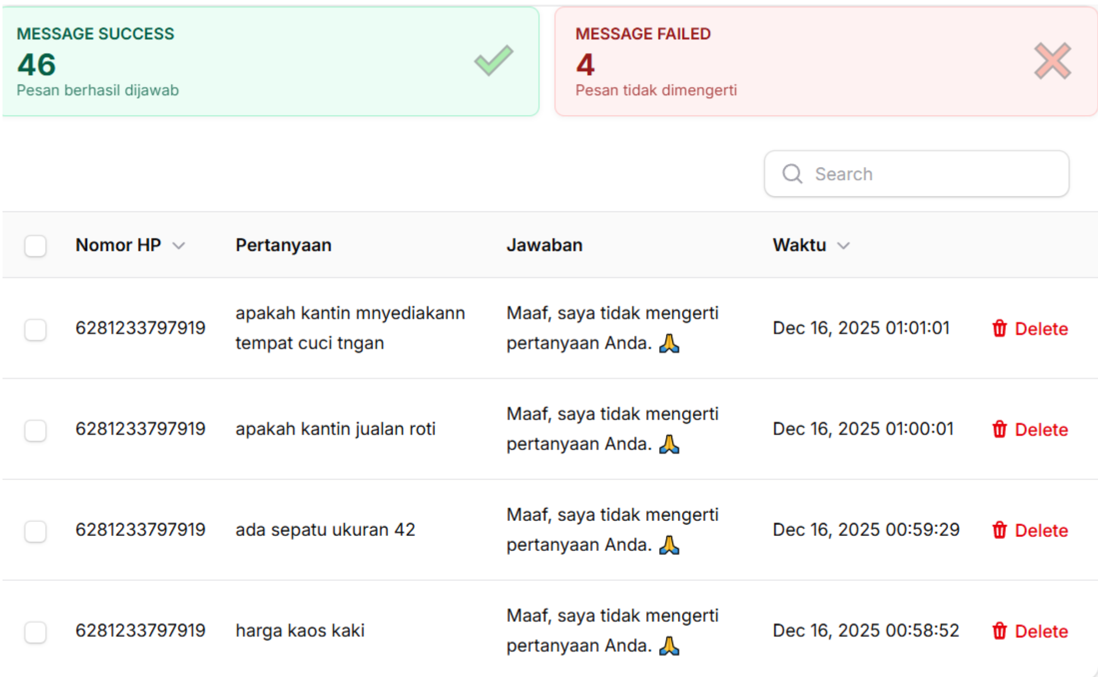

Figure 17. Message Failed view (4 unanswered)

3.8. Baseline comparison: SQL keyword query vs. cosine similarity

A baseline test was conducted to compare SQL keyword matching and Cosine Similarity for response retrieval. SQL performs well when queries contain exact stored keywords, but often fails on paraphrases or shorter phrasing. In contrast, Cosine Similarity is more robust to wording variation and produces more consistent matches. The comparison scenarios and outputs are summarized in Table 7.

Table 7. Comparison scenarios (SQL vs. cosine similarity)

Scenario	SQL Query result	Cosine Similarity result	Interpretation
"Jam buka kantin?"	Not found	"Canteen opens 08:00–16:00."	Cosine more flexible
"Menu apa hari ini?"	Correct	Correct	Both succeed
"Kantin buka hari minggu?"	Correct	Correct	Both succeed

4. Discussion

This study reinforces prior evidence that retrieval-based chatbots can improve service efficiency for repetitive, informational queries, especially when the system is designed around a structured FAQ/knowledge base rather than open-domain conversation. Earlier studies have reported that Cosine Similarity, often paired with TF–IDF, can return relevant answers quickly in chatbot settings such as student services and general information support (Attigeri et al., 2024; Setiawan & Adnyana, 2023; Shrivastava et al., 2022). Consistent with that line of work, the proposed system produced strong retrieval performance in a real operational context, evidenced by 92% accuracy (46/50) and a high F1-score (~0.95) under a fixed threshold policy (0.6). This supports the argument that vector-space similarity remains a practical baseline for controlled, routine domains.

At the same time, the contribution of this study is contextual: most prior chatbot implementations emphasize education, customer support, or broad organizational helpdesks, while fewer focus on internal workplace services with time-sensitive content (daily menus) and shift-based access constraints. In this setting, manual menu dissemination via noticeboards or verbal communication is structurally prone to delayed reach and uneven access (Breathnach et al., 2020; Taylor et al., 2021). The WhatsApp deployment aligns with Indonesian organizational communication patterns reported in recent studies, where WhatsApp is widely adopted and reduces friction by avoiding new app installations (Elciyar, 2025; Oyewale et al., 2026). Therefore, the system's relevance extends beyond the algorithmic match to include channel fit and operational alignment with shift-work routines.

The findings also triangulate the "value" of Cosine Similarity through multiple, complementary validations, rather than relying on a single metric. Functional readiness was confirmed via black-box testing across all core modules (authentication, CRUD operations for menu/schedule/FAQ/users, logging, and response handling). Computational correctness was strengthened by the manual-versus-system similarity verification (0.816 vs 0.8165), showing the implementation matches the expected mathematics. Retrieval robustness was then assessed with threshold tolerance on both correctly written and informal/misspelled queries, demonstrating that performance remains high but degrades for a small subset of noisy variants (4 unmatched). Finally, the baseline comparison against SQL/keyword matching provides converging evidence: keyword methods perform adequately when phrasing overlaps exactly, but Cosine Similarity is more tolerant of paraphrase and short phrasing, which aligns with patterns reported in the information retrieval literature.

Finally, the decision logic (Static/Menu/Schedule with fallback) strengthens the system's reliability in ways that many generic chatbot studies under-report. The explicit fallback behavior prevents the chatbot from answering out-of-domain questions (e.g., batik shirt prices), reducing the risk of misleading responses and aligning with best practices for constrained-domain conversational systems. The logging mechanism further connects system performance to continuous improvement: unmatched queries and emerging phrasing variants serve as empirical signals for dataset refinement, creating a maintainable loop in which the knowledge base evolves based on real usage rather than assumptions.

5. Limitations and Implications

This study has several limitations that should be considered when interpreting the results. First, it is a single-site case study conducted only at the employee canteen of Matahari Department Store Sunrise Mall Mojokerto, so the findings may not fully generalize to other workplaces with different communication cultures, menu dynamics, or employee language patterns. Second, the evaluation dataset was relatively small (50 test questions), which is useful for initial validation but may not fully capture the diversity of real-world queries over time, especially as menu items and employee phrasing evolve. Third, although Cosine Similarity with TF/TF–IDF performed well, the approach remains lexical-overlap-dependent, meaning it can still miss matches when employees use heavy abbreviations, slang, or uncommon paraphrases, as reflected in the 4 unmatched cases under the 0.6 threshold. Fourth, system performance is partly dependent on the reliability of the WhatsApp gateway (Wablas) and network stability; changes in platform policies or outages could affect end-to-end delivery even if the internal algorithm works correctly. Finally, this study emphasizes functional and retrieval metrics but does not yet provide long-term evidence on user satisfaction, consistent adoption, reduced staff workload, or sustained operational impact beyond the testing period.

The results imply that a WhatsApp-based, retrieval-oriented chatbot is a practical and scalable solution for improving internal information delivery in shift-based workplaces, particularly for routine and time-sensitive content such as daily menus and canteen schedules. With 92% accuracy and a high F1-score (~0.95), the system demonstrates that Cosine Similarity can serve as an effective baseline method for controlled domains where responses are drawn from a structured knowledge base, and it outperforms simple SQL/keyword matching when employee queries vary in phrasing.

Operationally, the separation of roles, employees accessing information via WhatsApp, and admins maintaining menus, schedules, FAQs, and logs via a web panel, supports governance, auditability, and continuous improvement. The logging feature is especially important because it creates a data-driven feedback loop: unmatched or new question variants can be identified and converted into additional FAQ patterns, gradually improving coverage. For future implementations, organizations can treat this workflow as a replicable model for other internal services (e.g., HR FAQs, shift schedules, facility information). In contrast, future research can extend the approach by testing larger, naturally collected query logs and comparing lexical similarity with more semantically oriented retrieval methods to reduce failures with informal language.

6. Conclusion and Future Directions

This study demonstrates that a retrieval-based WhatsApp chatbot can replace manual canteen announcements, providing a faster, more consistent information service in a shift-based workplace. By structuring the knowledge base into components such as the menu, schedule, and FAQs, and by utilizing preprocessing, TF/TF–IDF vectorization, and Cosine Similarity, the system effectively routed employee queries to the correct response type (Static/Menu/Schedule). This approach reduced the risk of incorrect responses, relying on a threshold-based fallback policy. Functional tests confirmed that the implementation, comprising a WhatsApp gateway to the application server and supported by an admin web panel, operated reliably. Empirical evaluation further demonstrated strong response relevance, with 92% accuracy and an F1 score of approximately 0.95, showcasing the system's resilience against paraphrasing compared to simpler SQL/keyword matching. This supports its practicality for addressing routine internal information needs.

Future work should focus on validating the system's generalizability by deploying the chatbot across multiple workplaces and evaluating its performance on larger, naturally collected query logs over extended periods, including during seasonal menu changes. To address unmatched cases caused by abbreviations, slang, and heavy paraphrasing, the matching module can be enhanced with semantic retrieval techniques, such as sentence embeddings, or a hybrid approach that combines lexical and semantic matching while maintaining the threshold safeguard. Additional evaluations should assess user satisfaction, adoption frequency, reductions in workload for canteen staff, and operational impact, such as a decrease in repeated inquiries. From an engineering standpoint, future iterations should prioritize enhancing reliability through gateway failover mechanisms, monitoring, and alerting systems, while also reinforcing privacy governance with retention policies, role-based access control, and audit trails. Furthermore, the introduction of analytics-driven dataset refinement workflows will ensure improvements are made systematically rather than ad hoc.

Author’s declaration

Author contribution

Ramadhani Handika Saputra: Conceptualization; Methodology; Software; Data curation; System design and implementation; Validation; Formal analysis; Writing original draft; Writing review & editing; Visualization. **Hadi Sucipto**: Conceptualization; Methodology; Data curation; Investigation; Supervision; Writing review & editing; Project administration.

Funding statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements

The author thanks the management and canteen staff of Matahari Department Store Sunrise Mall Mojokerto for granting permission to conduct the study and for supporting observation and interviews during the system development and evaluation.

Conflict of interest

The author declares no competing interests.

Ethical clearance

Research permission was obtained from the host institution (Matahari Department Store Sunrise Mall Mojokerto). Participation in interviews was voluntary and based on informed consent. This study was conducted in accordance with the principles of the Declaration of Helsinki. No sensitive personal data was collected; chat logs used for evaluation were minimized and anonymized.

AI statement

The author used a grammar-checking tool, namely Grammarly, to improve language clarity. The tool was not used to generate scientific content, and all text, tables, and figures were produced and verified manually by the author.

Publisher's and Journal's note

Universitas Negeri Padang as the publisher, and the Editor of Jurnal Pendidikan Teknologi Kejuruan state that there is no conflict of interest towards this article publication.

References

Adu-Gyamfi, M., He, Z., Nyame, G., Boahen, S., & Frempong, M. F. (2021). Effects of internal csr activities on social performance: the employee perspective. *Sustainability (Switzerland)*, *13*(11), 1–28. https://doi.org/10.3390/su13116235

Alateeg, S., & Alhammadi, A. (2024). The Impact of Organizational Culture on Organizational Innovation with mediation role of Strategic Leadership in Saudi Arabia. *Journal of Statistics Applications and Probability*, *13*(2), 843–858. https://doi.org/10.18576/jsap/130220

Amrit, C., & Narayanappa, A. K. (2025). An analysis of the challenges in the adoption of MLOps. *Journal of Innovation and Knowledge*, *10*(1), 100637. https://doi.org/10.1016/j.jik.2024.100637

Andrade, I. M. De, & Tumelero, C. (2022). Increasing customer service efficiency through artificial intelligence chatbot. *Revista de Gestao*, *29*(3), 238–251. https://doi.org/10.1108/REGE-07-2021-0120

Aristovnik, A., Ravšelj, D., & Murko, E. (2024). Decoding the Digital Landscape: An Empirically Validated Model for Assessing Digitalisation across Public Administration Levels. *Administrative Sciences*, *14*(3), 1–22. https://doi.org/10.3390/admsci14030041

Attigeri, G., Agrawal, A., & Kolekar, S. V. (2024). Advanced NLP Models for Technical University Information Chatbots: Development and Comparative Analysis. *IEEE Access*, *12*, 29633–29647. https://doi.org/10.1109/ACCESS.2024.3368382

Blagoev, B., Hernes, T., Kunisch, S., & Schultz, M. (2024). Time as a Research Lens: A Conceptual Review and Research Agenda. *Journal of Management*, *50*(6), 2152–2196. https://doi.org/10.1177/01492063231215032

Branda, F., Stella, M., Ceccarelli, C., Cabitza, F., Ceccarelli, G., Maruotti, A., Ciccozzi, M., & Scarpa, F. (2025). The Role of AI-Based Chatbots in Public Health Emergencies: A Narrative Review. *Future Internet*, *17*(4), 1–32. https://doi.org/10.3390/fi17040145

Breathnach, S., Llewellyn, C. H., Koutoukidis, D. A., van Rugge, C. R., Sutherland, A., & Lally, P. (2020). Experience of using an online pre-ordering system for a workplace canteen that offers lower-energy swaps: A think-aloud study. *Nutrients*, *12*(12), 1–26. https://doi.org/10.3390/nu12123878

Cedergren, A., & Hassel, H. (2024). Building organizational adaptive capacity in the face of crisis: Lessons from a public sector case study. *International Journal of Disaster Risk Reduction*, *100*, 104235. https://doi.org/10.1016/j.ijdrr.2023.104235

Chen, T., Gascó-Hernandez, M., & Esteve, M. (2024). The Adoption and Implementation of Artificial Intelligence Chatbots in Public Organizations: Evidence from U.S. State Governments. *American Review of Public Administration*, *54*(3), 255–270. https://doi.org/10.1177/02750740231200522

Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., Filieri, R., Jacobson, J., Jain, V., Karjaluoto, H., Kefi, H., Krishen, A. S., Kumar, V., Rahman, M. M., Raman, R., Rauschnabel, P. A., Rowley, J., Salo, J., Tran, G. A., & Wang, Y. (2021). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*, *59*, 102168. https://doi.org/10.1016/j.ijinfomgt.2020.102168

Elciyar, K. (2025). Users' privacy behaviors in response to WhatsApp policy changes. *Information and Computer Security*, *33*(4), 573–589. https://doi.org/10.1108/ICS-04-2024-0096

Hayati, C. (2019). Shopping Motivations and Their Influence on Shopping Experience in Sunrise Mall Mojokerto. *Media Mahardhika*, *17*(2), 264. https://doi.org/10.29062/mahardika.v17i2.83

He, W., Zhang, Z. (Justin), & Li, W. (2021). Information technology solutions, challenges, and suggestions for tackling the COVID-19 pandemic. *International Journal of Information Management*, *57*, 102287. https://doi.org/10.1016/j.ijinfomgt.2020.102287

Himpong, M. D., Waleleng, G. J., & Warouw, D. M. D. (2023). The Function of Whatsapp Group as a Medium of Communication Government and People in Sinisir Village South Minahasa Regency. *Journal La Sociale*, *4*(5), 301–110. https://doi.org/10.37899/journal-la-sociale.v4i5.900

Kernan Freire, S., He, T., Wang, C., Niforatos, E., & Bozzon, A. (2025). Factory operators' perspectives on cognitive assistants for knowledge sharing: challenges, risks, and impact on work. *Behaviour and Information Technology*, 1–27. https://doi.org/10.1080/0144929X.2025.2574373

Laymouna, M., Ma, Y., Lessard, D., Schuster, T., Engler, K., & Lebouché, B. (2024). Roles, Users, Benefits, and Limitations of Chatbots in Health Care: Rapid Review. *Journal of Medical Internet Research*, *26*, 1–28. https://doi.org/10.2196/56930

Liu, H., Chen, X., & Liu, X. (2022). A Study of the Application of Weight Distributing Method Combining Sentiment Dictionary and TF-IDF for Text Sentiment Analysis. *IEEE Access*, *10*, 32280–32289. https://doi.org/10.1109/ACCESS.2022.3160172

Lokman, A. S., Ameedeen, M. A., & Ghani, N. A. (2020). A Conceptual IR Chatbot Framework with Automated Keywords-based Vector Representation Generation. *IOP Conference Series: Materials Science and Engineering*, *769*, 1–9. https://doi.org/10.1088/1757-899X/769/1/012020

Moganadas, S. R., & Goh, G. G. G. (2022). Digital Employee Experience Constructs and Measurement Framework: A Review and Synthesis. *International Journal of Technology*, *13*(5), 999–1012. https://doi.org/10.14716/ijtech.v13i5.5830

Mohamad Suhaili, S., Salim, N., & Jambli, M. N. (2021). Service chatbots: A systematic review. *Expert Systems with Applications*, *184*(July), 115461. https://doi.org/10.1016/j.eswa.2021.115461

Nidhra, S. (2012). Black Box and White Box Testing Techniques - A Literature Review. *International Journal of Embedded Systems and Applications*, *2*(2), 29–50. https://doi.org/10.5121/ijesa.2012.2204

Oyewale, C. T., Osei, E., Gyasi, S. K., Boahene, U. O., & Adeyemi, F. A. (2026). A Descriptive Analysis of WhatsApp Usage: Media Sharing, Storage Practices, and Sustainability Implications. *Proceedings of Fifth International Conference on Computing and Communication Networks*, 329–343. https://doi.org/10.1007/978-3-032-14197-2_29

P, T. S., & G, K. (2024). Systematic review on utilization of WhatsApp in education. *E-Learning and Digital Media*, 1–14. https://doi.org/10.1177/20427530241239424

Paul, J., Ueno, A., Dennis, C., Alamanos, E., Curtis, L., Foroudi, P., Kacprzak, A., Kunz, W. H., Liu, J., Marvi, R., Nair, S. L. S., Ozdemir, O., Pantano, E., Papadopoulos, T., Petit, O., Tyagi, S., & Wirtz, J. (2024). Digital transformation: A multidisciplinary perspective and future research agenda. *International Journal of Consumer Studies*, *48*(2), 1–28. https://doi.org/10.1111/ijcs.13015

Petersen, K., Wohlin, C., & Baca, D. (2009). The Waterfall Model in Large-Scale Development. *Product-Focused Software Process Improvement*, 386–400. https://doi.org/10.1007/978-3-642-02152-7_29

Sachdeva, A., Kim, A., & Dennis, A. R. (2024). Taking the Chat out of Chatbot? Collecting User Reviews with Chatbots and Web Forms. *Journal of Management Information Systems*, *41*(1), 146–177. https://doi.org/10.1080/07421222.2023.2301175

Saravanos, A., & Curinga, M. X. (2023). Simulating the Software Development Lifecycle: The Waterfall Model. *Applied System Innovation*, *6*(6), 1–19. https://doi.org/10.3390/asi6060108

Setiawan, G. H., & Adnyana, I. M. B. (2023). Improving Helpdesk Chatbot Performance with Term Frequency-Inverse Document Frequency (TF-IDF) and Cosine Similarity Models. *Journal of Applied Informatics and Computing*, *7*(2), 252–257. https://doi.org/10.30871/jaic.v7i2.6527

Shrivastava, R., Pujahari, A., Singh, S. P., & Bhowmik, T. (2022). Efficient Question Answering in Chatbot Using TF-IDF and Cosine Similarity. *Innovations in Information and Communication Technologies*, 25–31. https://doi.org/10.1007/978-981-19-3796-5_3

Shrivastava, R., Pujahari, A., Singh, S. P., & Bhowmik, T. (2023). Efficient Question Answering in Chatbot Using TF-IDF and Cosine Similarity. *Innovations in Information and Communication Technologies*, 25–31. https://doi.org/10.1007/978-981-19-3796-5_3

Skuridin, A., & Wynn, M. (2024). Chatbot Design and Implementation: Towards an Operational Model for Chatbots. *Information (Switzerland)*, *15*(4), 1–28. https://doi.org/10.3390/info15040226

Sofyani, H., Riyadh, H. A., & Fahlevi, H. (2020). Improving service quality, accountability and transparency of local government: The intervening role of information technology governance. *Cogent Business and Management*, *7*(1), 1–20. https://doi.org/10.1080/23311975.2020.1735690

Taylor, T. K., Das, R., Mueller, K. L., Pransky, G. S., Harber, P., McLellan, R. K., Hartenbaum, N. P., Behrman, A. J., Roy, D. R., Blink, R. C., & Dreger, M. (2021). Safely Returning America to Work Part II Industry-Specific Guidance. *Journal of Occupational and Environmental Medicine*, *63*(6), E373–E391. https://doi.org/10.1097/JOM.0000000000002206

Wang, W. T., & Wu, S. Y. (2021). Knowledge management based on information technology in response to COVID-19 crisis. *Knowledge Management Research and Practice*, *19*(4), 468–474. https://doi.org/10.1080/14778238.2020.1860665

Zhang, L. (2025). Features extraction based on Naïve Bayes algorithm and TF-IDF for news classification. *Plos One*, *20*, 1–17. https://doi.org/10.1371/journal.pone.0327347

Zhang, W., Yoshida, T., & Tang, X. (2011). A comparative study of TF*IDF, LSI and multi-words for text classification. *Expert Systems with Applications*, *38*(3), 2758–2765. https://doi.org/10.1016/j.eswa.2010.08.066