Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

June 2026. Vol. 5. No. 3; e-ISSN: 2808-4519

AI-Based Chatbot Development for Academic Information Services at Universitas Negeri Medan and an Analysis of Indonesian Language Usage

Muhammad Rafli Wijaya[1]*, Alya Namira[2], Ananda Syafika[3], Syairal Fahmy Dalimunthe[4]

[1,2,3]Computer Science, Universitas Negeri Medan
[4]Indonesian Literary Studies, Universitas Negeri Medan
Jl. Willem Iskandar, Ps. V, Medan Estate, Kec. Percut Sei Tuan, Kab. Deli Serdang, Kota Medan – Sumatera Utara
rafliwijaya2005@gmail.com[1], alyanamira3010@gmail.com[2], anandasyafika402@gmail.com[3], fahmy@unimed.ac.id[4]*

Abstract

The efficiency of academic information services remains a major challenge in higher education institutions, particularly in responding to student inquiries quickly and independently. This study develops an artificial intelligence-based chatbot system for academic information services at Universitas Negeri Medan (UNIMED) by utilizing the LLaMA 3.1 8B Instant Large Language Model (LLM) via the Groq API within a Retrieval-Augmented Generation (RAG) framework. The system was built using a three-tier architecture consisting of a React.js (Vite) frontend, a Node.js with Express backend, and a Supabase (PostgreSQL) database serving as the academic FAQ knowledge base. The Cross-Industry Standard Process for Data Mining (CRISP-DM) was adopted as the research methodology. System evaluation was conducted using Black-Box Testing across four main scenarios: questions available in the FAQ, follow-up questions requiring conversational context, off-topic questions, and reference link validation all of which yielded a pass status. Furthermore, Indonesian language testing demonstrated that the system is capable of understanding diverse student language variations, including formal language, informal expressions, academic abbreviations, and ambiguous queries, while maintaining appropriate academic communication etiquette. The results indicate that the RAG approach is effective in reducing AI hallucination risks, and that this web-based chatbot offers broader accessibility compared to previous messaging platform-based chatbot systems.

***Keywords**: Information Academic Chatbot, Retrieval-Augmented Generation, Academic Information Service, Llama Chatbot, Artificial Intelligence Based Chatbot*

1. Introduction

In the era of educational digitalization, the efficiency of information services has become a primary benchmark for service quality in higher education institutions. This phenomenon is clearly visible in the creative industry and e-commerce sectors, such as Shopee or Tokopedia, where initial interactions between buyers and sellers are now dominated by artificial intelligence-based virtual assistants to provide instant responses [1]. The success of AI implementation in these commercial sectors demonstrates great potential if applied within the academic ecosystem, specifically on service portals such as Devakad or University Portals, to overcome the limitations of manual responses from administrative staff.

At present, students demand real-time and independent access to academic information. However, obstacles frequently arise on official university websites that still have limited information features, such as the unavailability of a comprehensive Frequently Asked Questions (FAQ) page. This leads to an accumulation of repetitive questions regarding basic procedures, such as the guidelines for filling out the Course Selection Sheet (KRS) or official link information for tuition (UKT) payments. Without an automation system, the information delivery process becomes inefficient and hinders students' academic mobility.

As a cutting-edge solution, the use of Large Language Models (LLMs) such as Meta's LLaMA offers far superior language processing capabilities compared to rule-based chatbots. The Natural Language Processing (NLP) technology in LLaMA allows the system to understand sentence context in the Indonesian language more naturally, even when students use non-standard language varieties or abbreviations [2]. This semantic understanding capability is crucial so that the chatbot can accurately recognize the user's intent without having to integrate directly into confidential internal databases.

In addition to the technical aspects, the use of the Indonesian language in academic chatbots must also consider sociolinguistic dimensions. As a service within an educational institution, the language generated by the AI must be able to balance a casual, communicative nature for easy understanding with upholding the norms of academic politeness. This is highly important so that the interaction between students and technology still feels human, trustworthy, and aligned with the communication ethics of a higher education environment [3].

In view of this urgency, this study focuses on the implementation of an Indonesian-language AI chatbot by utilizing the LLaMA model as the language processing engine for academic information services. The focus of the study is directed at how this model processes variations in students' language inputs regarding administrative obstacles and the extent to which the accuracy of the generated information can help bridge the information gap on academic portals. This research is expected to contribute to the development of communication models for virtual assistants that are smarter, more responsive, and more adaptive to the needs of students in Indonesia.

2. Theoretical Basis

2.1. Chatbot

A chatbot is a computer program designed to simulate conversation with human users, either through text or voice. In its application, a chatbot functions not only as a machine capable of communicating with humans, but also serves to assist and serve people in various fields such as education, e-commerce, and business by acting as an automated online assistant [1].

2.2. Large Language Model (LLM)

A Large Language Model (LLM) is a type of artificial intelligence (AI) trained on a massive amount of text data to understand, summarize, translate, and generate text that resembles human communication. According to research conducted by Fan et al., the field of LLMs is developing very rapidly, with new research and development continuously emerging; they also added that LLMs have demonstrated excellent performance on various NLP tasks [4].

However, despite its massive development, LLMs face a major obstacle in specialized use cases that require in-depth knowledge. The main constraint is 'hallucination,' a phenomenon where the model tends to present incorrect or illogical data with a highly convincing tone [5].

2.3. Retrieval-Augmented Generation (RAG)

To overcome these limitations, the Retrieval-Augmented Generation (RAG) framework is utilized. RAG consists of two main components: an information retrieval system and a neural language model. In RAG, data obtained from external sources is combined with the language model's existing knowledge base to build a more comprehensive context. This integration process significantly sharpens the accuracy and relevance of the content, as the model is able to leverage specific information and the latest data originating from external sources [1]. Several previous studies used the Retrieval-Augmented Generation (RAG) approach, a method that integrates an external database as the primary reference source for the language model in formulating answers [6]. Research conducted by Darmawan et al. [7]applied RAG for a faculty academic guide chatbot system, while research by Elysia & Herianto [1] developed RAG as a chatbot to improve school information services.

2.4. Academic Information System

RAG's ability to generate accurate and relevant answers based on official documents makes it a highly potential approach to be applied to academic information services. However, in reality, many educational institutions still rely on conventional methods to convey information to students, such as through one-way websites or manual services by administrators, which causes slow responses to user questions [8].

Several studies have attempted to address this issue by developing chatbots based on messaging platforms such as Telegram. Qalimaturrahmah & Santoso [8]developed an academic service chatbot using the RASA framework integrated into Telegram, successfully providing appropriate responses to student questions. Similarly, Ikhsan et al. [9] also built a campus information chatbot for Universitas Negeri Medan that operates through Telegram with an accuracy rate of around 70%. Although showing promising results, both systems have limitations in terms of accessibility, as users are required to have and use the Telegram application to access them.

Based on this gap, this research develops a web-based academic chatbot using React and Node.js that can be accessed directly through a browser without requiring additional applications, by utilizing the Groq API and LLaMA model within the RAG framework to produce accurate and contextual responses.

2.5. Human-Computer Interaction

The interaction between humans and digital systems has become an important aspect in the development of chatbots as academic information services. Chatbots serve as a form of communication transformation from manual to automated services, thereby partially replacing the human role in providing information to users. In this context, chatbots allow students to obtain information quickly without having to interact directly with administrative staff.

Research conducted by Bariah et al. [10] shows that the use of chatbots in academic services can help students obtain information more easily and is capable of responding to questions quickly and continuously 24 hours a day. This confirms that the presence of chatbots not only increases service efficiency but also forms new interaction patterns between humans and technology.

In addition, the quality of user experience is also an important factor in this interaction. A system that is responsive, easy to use, and capable of providing relevant answers will enhance user comfort in accessing digital services.

2.6. Sociolinguistics in Digital Communication

In the context of digital communication, the linguistic aspect becomes a crucial factor, especially in the interaction between students and chatbots. Students, as users, generally exhibit variations in language use, ranging from formal language to casual or non-standard language.

Research on the use of slang among university students shows that the current generation tends to use informal language in daily communication because it is considered more flexible and intimate. However, the use of such language in an academic context can reduce the level of formality and clarity of communication [11]. Therefore, the chatbot system needs to be able to understand various language variations without disregarding academic norms.

On the other hand, digital interaction also frequently gives rise to the phenomena of code-mixing and language interference. Academic communication in the digital space is often influenced by spoken language and foreign languages, thereby creating ambiguity in message delivery [12]. This indicates that chatbots must be designed to adapt to the language dynamics of users, while still maintaining the proper and correct use of the Indonesian language.

2.7. Tehcnology Acceptance Model

Technology acceptance by users is a key factor in the successful implementation of chatbots within academic environments. One commonly used approach is the Technology Acceptance Model (TAM), which explains that the level of technology utilization is influenced by perceived ease of use and perceived usefulness by the users.

The results of the literature review indicate that chatbots can improve service efficiency as well as user satisfaction because they can provide instant and accurate responses. Furthermore, the ease of access through various digital platforms is also a factor that encourages users to utilize this technology more widely [13]. Accordingly, academic chatbots must be designed with a simple and functional interface to ensure ease of use and to provide real benefits for students, thereby increasing the acceptance rate of the technology.

2.8. Social Impact of Chatbot Use in Education

The use of chatbots in education brings various social impacts, both positive and negative. On the positive side, chatbots can improve information accessibility, accelerate service processes, and reduce the workload of administrative staff.

Research by Putra et al. [14] regarding chatbot development in academic information systems shows that this technology can serve as a solution to slow service issues and the limitations of manual interaction, thereby making the information delivery process more effective and efficient.

Nevertheless, the use of chatbots also carries potential negative impacts, such as reduced direct human interaction and the possibility of misinterpreting information. Furthermore, dependence on technology may also increase if users rely too heavily on automated systems to obtain information.

3. Research Methods

The methodology employed in this study is the Cross-Industry Standard Process for Data Mining (CRISP-DM), which consists of six systematic stages. The entire research workflow is visualized in Figure 1.

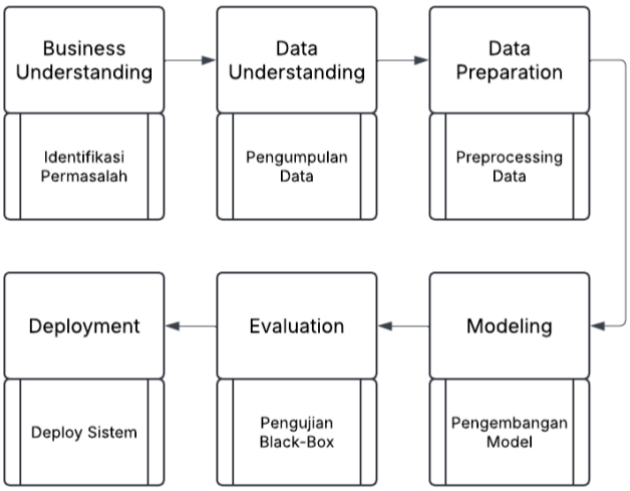

Figure 1: Research Workflow

3.1. Business Understanding

This stage was carried out by identifying the problems of academic information services at Universitas Negeri Medan (UNIMED), where students still face difficulties in obtaining academic information quickly and independently. Responses to student inquiries through conventional services were considered inefficient due to manual handling. Based on this condition, the proposed solution was the development of an artificial intelligence-based chatbot using the Retrieval-Augmented Generation (RAG) approach as a responsive and automated academic information service medium.

3.2. Data Understanding

The data used in this study were sourced from the official portal of Universitas Negeri Medan, including the university's official website (unimed.ac.id), the Devakad academic portal, the official guide page (panduan.unimed.ac.id), and UNIMED's official social media accounts. The collected data covers information regarding re-registration, KRS filling, UKT payment, SSO account creation, English Score guidelines, academic leave procedures, and other student-related information. Preliminary exploration was conducted to ensure the completeness and relevance of the data in relation to questions commonly asked by students.

3.3. Data Preparation

The collected data were processed into a structured FAQ format, where each entry consists of four main attributes: keywords (search keywords), question, answer, and link along with link_label as reference links to official documents. The total number of FAQ data entries produced amounted to 10–15 entries covering the most frequently asked academic topics. All FAQ data were subsequently stored in the Supabase database as the knowledge base of the chatbot system.

3.4. Modeling

The chatbot system was built using a three-tier architecture, consisting of a React.js (Vite) frontend, a Node.js with Express backend, and a Supabase (PostgreSQL) database. The language model used is LLaMA 3.1 8B Instant, accessed via the Groq API. The RAG approach was implemented with a relevance-score-based keyword matching mechanism to retrieve the FAQ entry most relevant to the user's question, which is then injected into the system prompt before being sent to the language model to generate an answer.

3.5. Evaluation

System testing was conducted using the Black-Box Testing method to verify overall system functionality from the end user's perspective without examining the internal code structure. The test scenarios covered four main conditions: questions available in the FAQ, follow-up questions requiring conversational history context, questions outside the academic topic (off-topic), and validation of reference links displayed. The system was declared successful if all scenarios produced responses consistent with the expected behavior.

3.6. Deployment

Upon meeting the evaluation criteria, the chatbot system was deployed online by separating the frontend and backend components. The frontend was deployed using Vercel, while the backend was deployed using Railway. The FAQ database uses Supabase, which can be updated directly without modifying the system code, thereby facilitating maintenance and the addition of academic information data in the future.

4. Results and Discussion

4.1. System Implementation

The implementation of this academic information service chatbot system was built using a three-tier architecture that dynamically integrates frontend, backend, and database components. The technology specifications used in the implementation of this academic chatbot are summarized in detail in the following table:

Component	Technology
Frontend	React.js (Vite)
Backend	Node.js + Express
Database	Supabase
Model Bahasa	LlaMa 3.1 8B Instant
Inferensi	Groq API
Deploymeny	Vercel + Railway

Table 1: Technology Specifications Used

4.1.1. System Interface

On the user side, the web interface was designed using React.js with the Vite build tool to ensure fast, lightweight, and responsive performance when accessed directly through the student's browser.

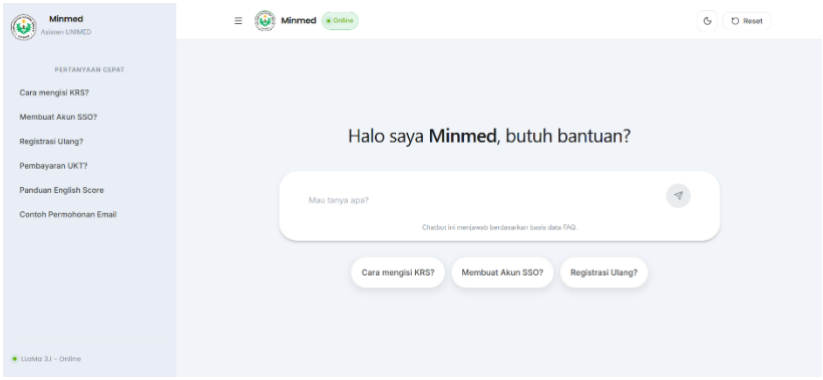

Figure 2: Welcome screen main page display

The welcome screen is the main landing page that students first encounter when opening the chatbot website. This page presents an interactive greeting while also providing a brief menu of academic topics that users can inquire about.

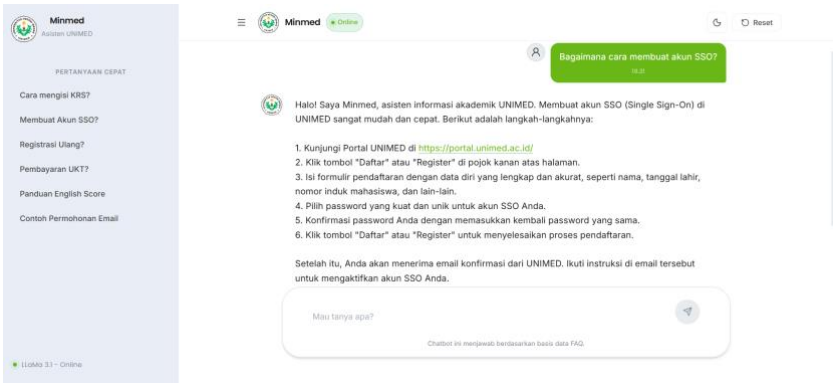

Figure 3: Normal conversation display

The normal conversation display shows real-time two-way interaction when students pose questions regarding administrative or academic issues. The system is capable of recognizing user intent and processing non-standard language input variations naturally, owing to the implementation of the LLaMA model.

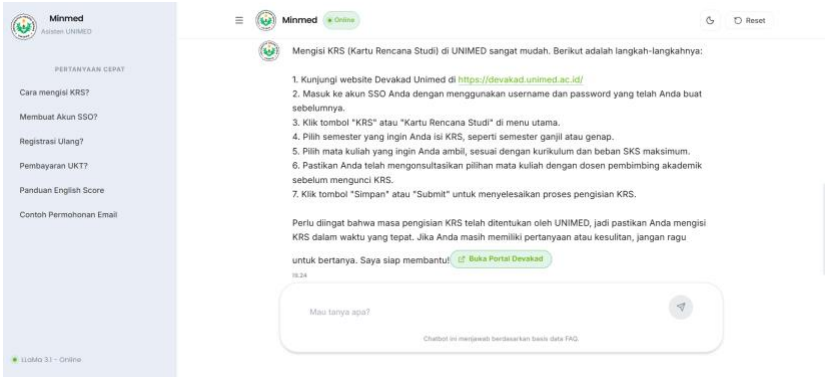

Figure 4: Reference link button display

This display highlights an important feature: the inclusion of a reference link button at the end of answers generated by the chatbot. These link buttons automatically direct students to official university documents or portal pages, such as Devakad or the UNIMED guide site, for further information validation. The presence of this external-data-based reference feature is critical in sharpening information accuracy while mitigating the potential "hallucination" effects of the AI model.

4.2. Implementation of RAG and FAQ Database

The implementation of this system centers on the provision of a reliable knowledge base to suppress the potential for data hallucination in the AI model. Academic data collected from various official portals of Universitas Negeri Medan were processed into a structured Frequently Asked Questions (FAQ) database format. All data were stored in the Supabase database using PostgreSQL, with a table structure specifically designed to facilitate keyword-matching-based search. The specifications and structure of the FAQ table in Supabase are described in detail in the following table:

Column	Data Type	Description
Id	SERIAL	Primary key
keywords	TEXT[]	Search keywords
question	TEXT	Question
answer	TEXT	Answer
link	TEXT	Reference URL
link_label	TEXT	Link button label

Table 2: Faq Table Structure

To provide a clear picture of the implemented data structure, the following are sample FAQ data entries representing academic topics such as KRS (Course Selection Sheet) filling and UKT (Single Tuition Fee) payment:

id	keywords	question	answer	link	link_label
1	["krs", "kartu rencana studi", "pengisian krs", "mata kuliah"]	How to fill out the KRS?	KRS is filled out via the UNIMED Devakad website during the designated KRS filling period. Students can choose courses according to the curriculum and maximum credit load. Consult course selections with your academic advisor before locking the KRS.	https://devakad.unimed.ac.id	Open Devakad Portal
2	["ukt", "uang kuliah tunggal", "biaya kuliah", "pembayaran kuliah", "spp"]	How to pay UKT?	UKT (Uang Kuliah Tunggal / Single Tuition Fee) is paid every semester through university-designated banks (BNI/BRI/Mandiri/BSI). Students can print the billing statement via UNIMED E-Billing.	https://panduan.unimed.ac.id/lihat_data_home/39	Open Guide

Table 3: Sample Faq Data Entries In Supabase

The Retrieval-Augmented Generation (RAG) workflow begins when the user submits a question through the chatbot interface. The question is forwarded to the Node.js backend, where the system executes a relevance-score-based keyword matching mechanism to match the user's input keywords against the keywords column available in the Supabase database. Once the most relevant FAQ entry is identified and retrieved from the database, the content — comprising the question and answer text — is directly injected into the system prompt as the primary reference context. This enriched prompt is then transmitted to the LLaMA 3.1 8B Instant model via the Groq API, enabling the model to generate a final answer that is accurate, contextual, and fully grounded in official internal documents.

4.3. System Testing Results

Testing of this chatbot system was conducted using the Black-Box Testing method to verify the overall functionality and performance of the system from the end user's perspective, without examining or testing the internal code structure. This evaluation focused entirely on interface testing to ensure that every variation of text input submitted by students can produce an appropriate, relevant, and functional response output in accordance with the system design. This testing is particularly important to validate the performance of the Retrieval-Augmented Generation (RAG) framework in processing documents, maintaining conversational context, and filtering requests outside the

academic domain. A summary of the system testing results using the Black-Box Testing method is presented in detail in the following table.

No	Test Scenario	Input	Expected Output	Status
1	Question available in FAQ	"Cara membayar UKT?"	System provides a valid FAQ answer regarding the KRS filling procedure on Devakad, along with a reference link button.	Pass
2	Follow-up question	"Terus setelah itu apa?"	System understands the previous conversation history and explains the next steps after filling out the KRS.	Pass
3	Off-topic question	"Bisa bantu coding?"	System recognizes topic limitations and declines to answer with an academically polite rejection message.	Pass
4	Reference link validation	Click "Buka Panduan" button	System opens a new browser tab directly navigating to the valid guide URL without a broken link.	Pass

Table 4: Chatbot System Testing Results

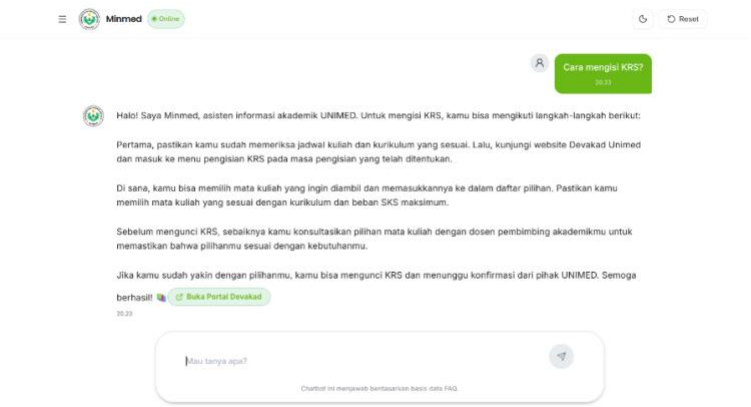

Figure 5: Testing Scenario 1 (Question in FAQ)

In the first scenario, testing was performed by directly providing input in the form of a common administrative question asked by students. The figure above shows that the chatbot successfully matched the input keywords with the Supabase knowledge base, extracted the relevant data, and assembled it into an informative answer complete with a reference link button.

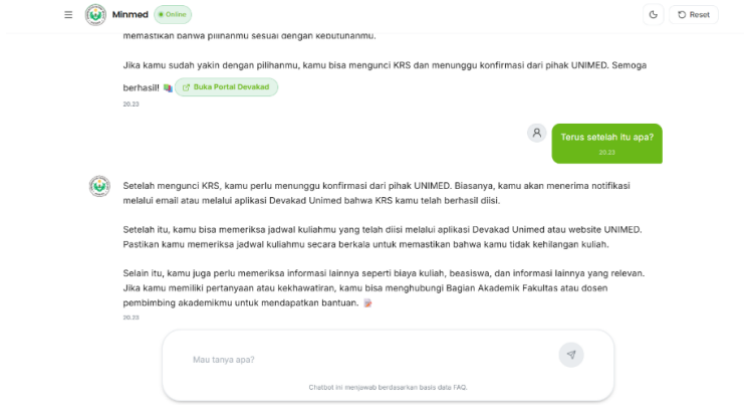

Figure 6: Testing Scenario 2 (Follow-up Question)

The second scenario aimed to test the LLaMA 3.1 model's ability to maintain contextual memory through the backend's conversation history parameter. Based on the testing results in the figure, when the user posed an ambiguous question such as "Then what should I do after that?", the system proved capable of correctly linking the continuation of the KRS filling procedure from the previous message, without confusion.

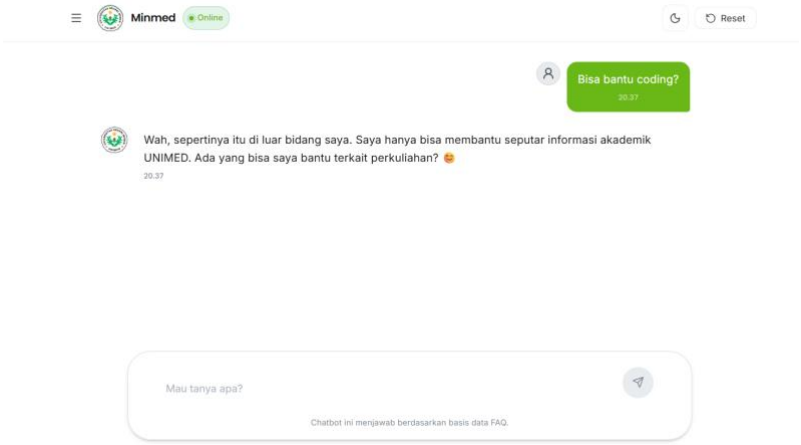

Figure 7: Testing Scenario 3 (Off-Topic Question)

The third scenario was designed as a form of scope restriction for the AI to prevent misuse for non-academic purposes on campus. Testing demonstrated the system's success in detecting off-topic inputs (such as requests to write programming code), with the chatbot immediately responding with a communicative rejection statement while maintaining formal academic etiquette.

Figure 8: Testing Scenario 4 (Reference Link Validation to Devakad)

The final scenario was conducted to test the external link navigation functionality from the Supabase database attributes. Testing results showed that when the user clicked the reference action button displayed at the end of the chat, the system responsively opened a new browser window and accurately directed it to the targeted educational institution portal URL.

4.4. Results of the Testing on Indonesian Language Usage

Indonesian language testing was conducted to evaluate the chatbot's ability to comprehend language variations used by students in academic information services. This testing focused on the system's capability to recognize formal language, non-standard language, abbreviations, and to maintain language politeness in generated responses.

Testing was performed using several variations of student question inputs commonly used in everyday communication. The system was then observed based on the relevance of the answers, the ability to understand the intent of the questions, and the quality of Indonesian language usage in the chatbot's responses.

The testing results showed that the LLaMA 3.1 model used in the system is capable of understanding various student language variations quite well. The system consistently produced relevant answers even when users employed abbreviations, casual language, or non-formal sentence structures.

No	Language Test Scenario	Input	Expected Output	Status
1	Formal language	"Bagaimana Cara Pembayaran UKT?"	System provides a clear and formal academic answer regarding the UKT payment procedure.	Pass
2	Non-standard language	"Bayar UKT gimana ya bang?"	System still understands the intent of the question and provides a relevant answer regarding UKT payment.	Pass
3	Use of abbreviations	"Cara isi KRS gimana?"	System recognizes the term 'KRS' and correctly provides the KRS filling guide.	Pass
4	Ambiguous question	"Terus abis itu?"	System understands previous conversation context and provides an appropriate follow-up response.	Pass
5	Language etiquette	"Tolong bantu cara login Devakad"	System provides a polite, communicative, and easy-to-understand response.	Pass

Table 5: Results Of Indonesian Language Usage Testing

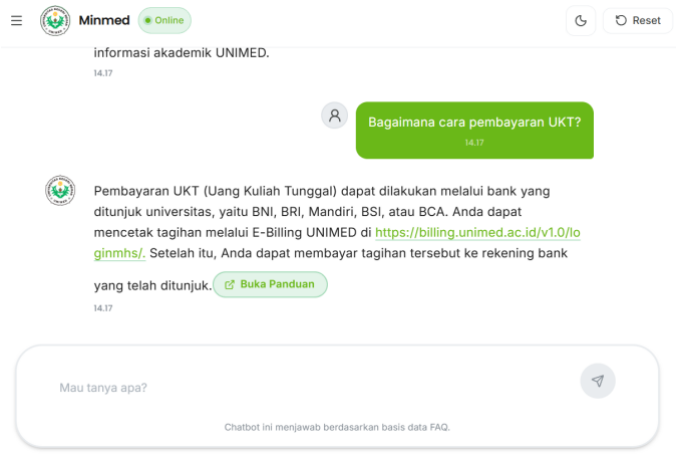

Figure 9: Language Testing Scenario 1 (Formal Language)

Based on the testing results in the figure above, the chatbot successfully comprehended the question conveyed using formal Indonesian, specifically regarding the UKT payment procedure at Universitas Negeri Medan. The system was able to provide a clear, informative response appropriate to the academic context, along with an official reference link as an additional information source. This demonstrates that the implementation of the LLaMA 3.1 model and the Retrieval-Augmented Generation (RAG) approach is capable of producing well-formed Indonesian language-based communication that is easily understood by users.

Figure 10: Language Testing Scenario 2 (Non-Standard Language)

In this test, the user submitted a question using informal Indonesian commonly found in students' everyday conversations, such as the use of "dimana ya bang" ("Where you paying, bro?"). Despite the non-standard input structure, the chatbot was still able to understand the intent of the question and produce a relevant response regarding UKT payment at Universitas Negeri Medan. This result demonstrates that the LLaMA 3.1 model possesses strong Natural Language Processing (NLP) capabilities in understanding informal language variations, making interactions feel more flexible and natural for users.

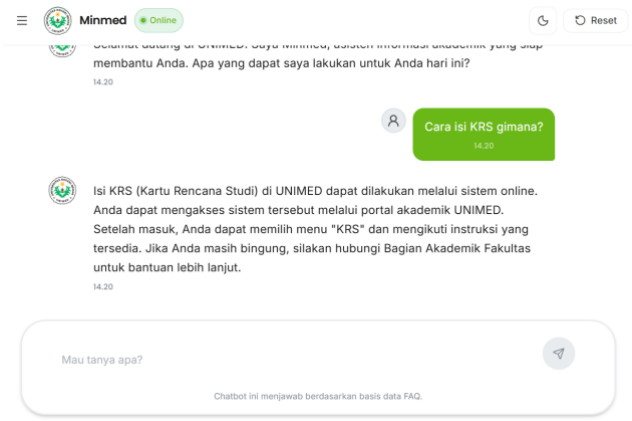

Figure 11: Language Testing Scenario 3 (Abbreviation Usage)

In this test, the user submitted a question using academic abbreviations and a non-formal sentence structure: "Cara isi KRS gimana?" ("How do I fill in the KRS?"). The testing results showed that the chatbot was able to recognize the term "KRS" as Kartu Rencana Studi (Course Selection Sheet) and correctly understand the user's intent. The system then provided an informative answer clearly and communicatively explaining the KRS filling procedure. This capability demonstrates that the LLaMA 3.1 language model is capable of comprehending common academic terms frequently used by students in daily communication.

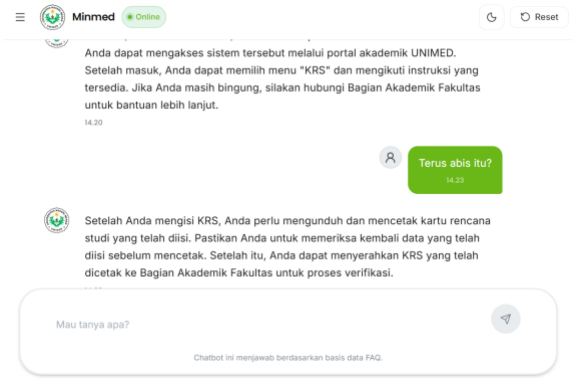

Figure 12: Language Testing Scenario 4 (Ambiguous Question)

In this test, the user submitted a follow-up question, "Terus abis itu?" ("Then what?"), after previously discussing the KRS filling process. Despite the ambiguous nature of the question, with no direct reference to a topic, the chatbot was still capable of understanding the previous conversational context and providing an appropriate follow-up response. This result demonstrates that the system possesses the ability to maintain conversation history, allowing the interaction between the user and the chatbot to proceed in a more natural, coherent, and communicative manner.

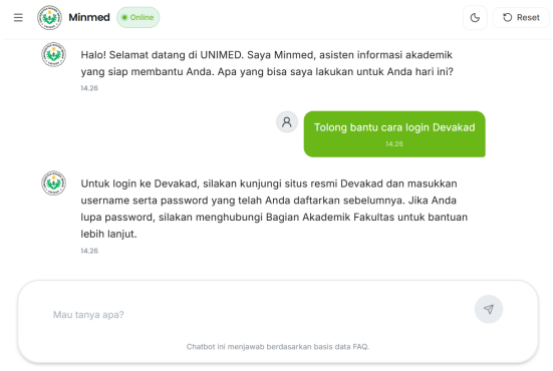

Figure 13: Language Testing Scenario 5 (Language Etiquette)

In this test, the user submitted a question in a polite manner: "Tolong bantu cara login Devakad" ("Please help me with how to log in to Devakad"). The testing results showed that the chatbot was capable of providing a communicative, formal response while maintaining language etiquette appropriate to the academic environment. In addition to clearly providing the login steps, the system also employed word choices that were easy to understand and professional. This demonstrates that the chatbot is not only capable of understanding user intent, but can also uphold academic communication etiquette in Indonesian language-based interactions.

4.5. Discussion

Based on the implementation and testing results that have been conducted, the RAG-based UNIMED academic information service chatbot system demonstrates adequate performance in meeting students' information needs. The following discussion focuses on three main aspects: the effectiveness of the RAG approach, system limitations, and a comparison with previous studies.

4.5.1. Effectiveness of the RAG Approach

The Black-Box Testing results show that all four test scenarios were successfully executed in accordance with the expected outputs. The relevance-score-based keyword matching mechanism proved effective in identifying the appropriate FAQ entries and injecting them as context into the LLaMA 3.1 model prompt. This approach directly suppresses the risk of hallucination, which is a major weakness of LLMs, as pointed out by Husain et al. [5] that models tend to present incorrect data with a convincing tone when not accompanied by valid external context. Furthermore, the implemented conversation history feature allows the system to maintain context between conversation turns, ensuring that follow-up questions from students can be answered coherently without losing the previous flow of information.

4.5.2. System Limitations

Although the testing results show success across all scenarios, several limitations must be acknowledged. First, the keyword matching mechanism used relies on exact string matching, meaning the system is not yet capable of capturing synonyms, abbreviations, or spelling variations that are not registered in the database's keywords column. Second, the FAQ database, which contains 10–15 entries, is still relatively limited in scope, so questions outside the registered topics cannot be answered substantively. Third, limiting the model's response length (max_tokens) potentially shortens answers when the transmitted conversation history context is quite long, resulting in less comprehensive responses to follow-up questions.

4.5.3. Comparison with Previous Studies

Compared to the study by Ikhsan et al. [9], which also built a Telegram-based UNIMED information chatbot using a Neural Network architecture and achieved an accuracy rate of around 70%, the system developed in this study offers broader accessibility as it can be accessed directly through a browser without requiring additional applications. Meanwhile, Bustomy et al. [6] utilized a ChromaDB vector database with semantic search for their RAG implementation, which is theoretically superior in capturing contextual meaning compared to the keyword matching used in this study. This indicates that the approach applied in this research can still be further developed by adopting embedding-based semantic search techniques to improve retrieval coverage and accuracy.

4.5.4. Analysis of Indonesian Language Usage

The use of the LLaMA 3.1 model in the chatbot system has proven to assist the understanding process of the Indonesian language in a more natural manner compared to rule-based chatbots. The system is capable of comprehending variations in student language styles, including the use of abbreviations, casual language, and questions that are not formally structured. This capability is achieved because the LLM model employs a Natural Language Processing (NLP) approach that understands the semantic context of sentences, rather than just direct word matching.

Furthermore, the language quality of the chatbot's responses also demonstrates the characteristics of polite and informative academic communication. This is vital because the chatbot is utilized within an educational environment, meaning the generated responses must not only be accurate but also maintain the formal communication etiquette between the institution and the students.

5. Conclusion

This research has successfully developed an artificial intelligence-based chatbot system for the academic information services of Universitas Negeri Medan by utilizing the LLaMA 3.1 8B Instruct model within a Retrieval-Augmented Generation (RAG) framework. Built upon a web-based three-tier architecture, the system is proven capable of delivering accurate and contextual responses to various student academic inquiries, while offering broader accessibility compared to previous chatbot systems that relied on messaging platforms like Telegram.

From a functional standpoint, all Black-Box Testing scenarios were successfully executed in accordance with the expected outputs, proving that the relevance-score-based keyword matching mechanism is effective in retrieving the appropriate FAQ entries and injecting them as context into the language model. Consequently, the risk of hallucination a primary weakness of LLMs is significantly suppressed.

The system also successfully maintains conversation history, ensuring that follow-up questions from students can be answered coherently without losing the previous flow of information.

Linguistically, the implementation of the LLaMA 3.1 model has proven capable of understanding various language variations used by students in daily communication, ranging from formal language, non-standard terms, and academic abbreviations, to ambiguous questions, while consistently generating polite responses aligned with academic communication etiquette. This indicates that LLM models possess natural language processing capabilities that are far more adaptive than conventional rule-based chatbot approaches.

Nevertheless, this system still has a number of limitations that need to be addressed in future development. The exact string-matching nature of the keyword matching mechanism is not yet able to capture synonyms or variations in spelling that are not registered in the database, the FAQ coverage remains limited to 10–15 entries causing the system to be unable to answer off-topic questions substantively, and token length limitations potentially result in less comprehensive responses during long conversations. Therefore, further development utilizing embedding-based semantic search is highly recommended to improve the overall coverage and accuracy of the system.

References

[1] S. Elysia and Herianto, "Chatbot Berbasis Retrieval Augmented Generation (RAG) untuk Peningkatan Layanan Informasi Sekolah," *Journal TIFDA (Technology Information and Data Analytic)*, vol. 1, no. 2, pp. 52–58, Dec. 2024, doi: 10.70491/tifda.v1i2.52.

[2] M. Abid Nadzif, Saefurrohman, and R. Soelistijadi, "Penggunaan Teknologi Natural Language Processing dalam Sistem Chatbot Untuk Peningkatan Layanan Informasi Administrasi Publik," *Indonesian Journal of Computer Science Attribution*, vol. 13, no. 1, pp. 2024–1227, 2024.

[3] M. Erline and Y. Christian, "A R T I C L E I N F O Web-Based Chatbot with Natural Language Processing and Knuth-Morris-Pratt (Case Study: Universitas Internasional Batam)," *Jurnal Sains dan Teknologi*, vol. 11, no. 1, pp. 132–141, 2022, doi: 10.23887/jst-undiksha.v11i1.

[4] L. Fan, L. Li, Z. Ma, S. Lee, H. Yu, and L. Hemphill, "A Bibliometric Review of Large Language Models Research from 2017 to 2023," 2024. doi: https://doi.org/10.1145/3664930.

[5] M. L. Husain, Y. Wibisono, and A. Anisyah, "Development of an Academic Services Chatbot Based on Retrieval-Augmented Generation (RAG)," *Brilliance: Research of Artificial Intelligence*, vol. 5, no. 2, pp. 727–735, Aug. 2025, doi: 10.47709/brilliance.v5i2.6719.

[6] M. Saddam Heykal Bustomy, M. Alfan Rosid, and H. Setiawan, "Implementasi Chatbot Informasi Akademik Berbasis Retrieval-Augmented Generation (RAG) Menggunakan LLaMA 3.1 [Implementation of an Academic Information Chatbot Based on Retrieval-Augmented Generation (RAG) Using LLaMA 3.1]," 2026. doi: https://doi.org/10.21070/ups.9995.

[7] F. Darmawan, W. Gusdya Purnama, and A. A. Nurcahyo, "Prototipe Sistem Chatbot Panduan Akademik Fakultas Teknik Unpas menggunakan Large Language Model," *Jurnal Sistem dan Informatika*, vol. 19, no. 2, 2025.

[8] M. Qalimaturrahmah and D. B. Santoso, "Aplikasi Layanan dan Informasi Akademik Berbasis Chatbot Telegram Menggunakan Natural Language Processing," *Jurnal Teknologi Informasi dan Komunikasi)*, vol. 8, no. 2, p. 2024, 2024, doi: 10.35870/jti.

[9] M Ikhsan, Dila Marta Putri, Siti Nurjanah, Asde Rahmawati, Fahrizal Fahrizal, and Bastul Wajhi Akramunnas, "Implementasi Teknologi Chatbot sebagai Media Informasi di Universitas Negeri Medan," *Jurnal Teknik Mesin, Industri, Elektro dan Informatika*, vol. 4, no. 1, pp. 265–277, Feb. 2025, doi: 10.55606/jtmei.v4i1.4820.

[10] S. H. Bariah #1, W. P. #2, and K. A. N. Imania, "Pengembangan Virtual Assistant Chatbot Berbasis Whatsapp Pada Pusat Layanan Informasi Mahasiswa Institut Pendidikan Indonesia-Garut," *Jurnal PETIK*, vol. 8, 2022.

[11] D. Melinda Br Bangun *et al.*, "PT. Media Akademik Publisher DOMINASI BAHASA GAUL DI KALANGAN GEN Z DALAM KONTEKS PRESENTASI AKADEMIK: STUDI DISKRIPTIF PADA MAHASISWA UNIVERSITAS TRUNOJOYO MADURA Aisyah Dwi Anggraini 5 Septian Trio Bagus S 6," *JMA)*, vol. 2, pp. 3031–5220, 2024, doi: 10.62281.

[12] S. A. Putri, Z. Alpris Wulandari, N. F. Manalu, and A. N. Wulandari, "Interaksi Akademik di Ruang Digital: Analisis Kontak Bahasa dalam Chat WhatsApp Mahasiswa Sastra Indonesia Universitas Negeri Medan," *Journal of Innovative Research*, vol. 03, pp. 188–193, 2026, [Online]. Available: https://ziaresearch.or.id/index.php/mesada

[13] M. R. Febriansyah *et al.*, "IMPLEMENTASI CHATBOT SEBAGAI VIRTUAL ASSISTANT : SYSTEMATIC LITERATURE REVIEW," 2025.

[14] G. Putu Mahendra Putra, A. Tenriawaru, P. Studi Ilmu Komputer, F. Matematika dan Ilmu Pengetahuan Alam, and U. Halu Oleo, "Rancang Bangun Virtual Assistant Chatbot Menggunakan Node.Js pada Layanan Sistem Informasi Akademik," 2023.