Tracer Study Design with WhatsApp Chatbot Integration Using Natural Language Processing

Wigananda Firdaus Putra Aditya

[a] Magister Teknologi Informasi UPN "Veteran" Jawa Timur, Jl.Raya Rungkut Madya, Gunung Anyar, Surabaya 60294,Indonesia
[1] wiganandafirdaus@gmail.com

ARTICLE INFO

Keywords
Whatsapp, NLP, Chatbot, Tracer Studies

ABSTRACT

Tracer studies are methods used by educational institutions to track alumni and evaluate the effectiveness of the education they provide. One of the main challenges in conducting these studies is the low response rate, often caused by lengthy surveys and a lack of interactive engagement. To address this issue, a WhatsApp chatbot system powered by Natural Language Processing (NLP) was developed. This system offers an interactive and user-friendly survey experience, enabling respondents to complete the survey directly via WhatsApp without needing to visit a website. The responses are automatically stored in Google Sheets through an API. Utilizing a microservices architecture, the project efficiently separates key components such as the WhatsApp API, NLP services, and the Google Sheets API—resulting in improved data collection efficiency and a more convenient survey process for respondents.

1. Introduction

A tracer study is a research activity conducted by higher education institutions to track and analyze the outcomes of their graduates [1]. It is considered essential because it serves as a tool to evaluate the performance of higher education institutions. Moreover, it has become one of the accreditation requirements mandated by the National Accreditation Board for Higher Education (BAN-PT), as part of the Self-Evaluation document required when submitting proposals to the Ministry of Education and Culture [2].

Based on previous experience and observations, the implementation of tracer studies faces several challenges. One of the main issues is the large number of survey questions, which often discourages alumni from completing the survey. Additionally, the current method of survey distribution is typically one-way—conducted through a static website with no interactive features—leading to disengagement and incomplete responses, ultimately reducing the overall response rate.

WhatsApp, as one of the most widely used social media platforms today, serves as a popular medium for both personal and group communication [3]. A chatbot, on the other hand, is a dialog-based system that processes sentences using keywords or phrases to retrieve relevant information [4]. The chatbot matches user input with predefined patterns stored in its knowledge base, which is derived from various sources [5].

Natural Language Processing (NLP) is a subfield of computer science and computational linguistics focused on enabling machines to understand and respond to human language [6][7]. Meanwhile, an Application Programming Interface (API) refers to a set of functions that allows different programs to communicate with each other. APIs act as connectors that integrate multiple applications across diverse platforms, often referred to as public APIs [8].

Given the aforementioned problems and references, a WhatsApp-based chatbot system powered by NLP technology can be developed to enhance tracer study implementation. This chatbot system automates data collection by facilitating interactive conversations with respondents, eliminating the need to access a website. Furthermore, it is integrated with the Google Sheets API, allowing survey responses to be stored automatically.

This project adopts a microservices architecture to separate critical components such as the WhatsApp API, NLP services, and Google Sheets API. This modular structure enhances flexibility, scalability, and ease of maintenance. The proposed system is expected to improve the efficiency of tracer studies by minimizing errors, accelerating data collection, and offering a more user-friendly experience for respondents.

2. Method

2.1. *Microservice* Approach

Microservices are a collection of small, independent processes that communicate with each other. others to form complex applications that are agnostic to any API language [9]. Microservices are used to break down applications into smaller, independent components, where each service is responsible for a specific function within the system. In this project, the main services used are:

- WhatsApp API Service: This service is responsible for receiving and sending messages via WhatsApp, which will be used by respondents to fill out the survey.
- NLP Service: This service processes text input sent by respondents via WhatsApp. Natural Language Processing (NLP) is used to understand the intent of the text sent by respondents and send relevant responses.
- Google Sheets API: This service is used to store survey response data from respondents. The responses can then be further processed by the admin.

2.2. Agile SDLC Methods

The agile method is a short-term software development model. Then, it requires rapid adaptation in dealing with every change [10]. Meanwhile, there is an opinion that Agile Development is a rapid software development method with the condition of changing needs that occur in a relatively short time[11] . In Agile, projects are divided into several iterations or sprints, each of which produces a functional and testable part of the system. With this approach, developers can release system features such as chatbot integration, NLP processing, and data storage in the Google Sheets API in stages.

2.2.1. Flow of Method Usage

In its implementation, this agile method will be carried out with the following stages:

- Initial Planning
  Develop a WhatsApp chatbot that is capable of collecting survey answers from respondents and using NLP to process the data, as well as saving the survey results into Google Sheets. Key features that need to be developed include:
  - WhatsApp API integration.
  - Survey data collection via chatbot.
  - Data storage using Google Sheets API.
  - Input processing with NLP to understand and analyze answers respondents.
- Sprint Planning
  - Sprint 1: Chatbot Development
    Focus on developing a chatbot that is connected to the WhatsApp API and is capable of greeting and sending basic survey questions.

- Sprint 2: Google Sheets API Integration
  Connecting a chatbot with Google Sheets to store survey responses and Ensure that each respondent's answer is automatically saved in Google Sheets.
- Sprint 3: NLP Implementation
  Adding NLP features to analyze user responses, such as answer classification or pattern recognition in survey data
- Sprint 4: Testing and Debugging
  Full system testing, including chatbot, Google Sheets, and NLP integration, Fixed bugs that appeared

- Development and Implementation
  - Feature Development
    Features are developed one by one in separate sprints. At the end of each sprint, the completed features are tested and evaluated.
  - Testing
    Testing is done at the end of each sprint to ensure that each module is working properly and as expected.
- Release
  - Once the main features such as data collection and storage are ready, the first release is made.
  - In the second release, NLP features were added to improve interaction and understanding of survey responses.

3. Results and Discussion

3.1. Use Case Diagram

In this project, the Use Case Diagram plays a crucial role in visualizing how various components interact within the chatbot-based tracer study system. It illustrates the relationships and communication flow between the Tracer Study Administrator, Respondents, the Natural Language Processing (NLP) module, and the Google Sheets API. Specifically, the diagram outlines how the process begins with the administrator configuring the survey and initiating communication through the WhatsApp chatbot. Respondents then receive and interact with the chatbot, which dynamically delivers questions and captures responses in real-time using NLP techniques.

Once the responses are collected, the system automatically sends the data to the Google Sheets API, where each entry is securely stored for further analysis. By mapping these interactions, the Use Case Diagram provides a clear overview of the end-to-end workflow—starting from the distribution of survey questions to the completion and storage of survey responses. This visualization not only facilitates better understanding of the system's operational flow but also helps stakeholders identify functional requirements, potential bottlenecks, and areas for further enhancement in the system's design and implementation..

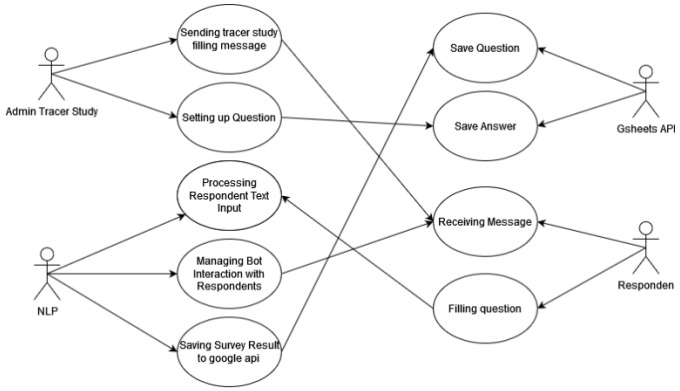

Figure 1. Use Case diagram of survey with WhatsApp bot

This Use Case Diagram illustrates the interaction flow in a chatbot-based tracer study system, where there are four main actors: Tracer Study Admin, Respondents, NLP, and Google Sheets API. Tracer Study Admin is responsible for sending survey completion messages and organizing questions that are then sent to Respondents via chatbot. Respondents play a role in answering questions given by the chatbot, and their answers are processed by NLP to understand the text automatically. The answers given by respondents are then stored by the Google Sheets API into Google Sheets. This diagram shows the relationship between these actors, from sending messages, filling out surveys, to storing the results, showing the entire interaction process in the system.

3.2. Class Diagram

The Class Diagram is designed to describe the overall structure and relationships among the various classes involved in the WhatsApp chatbot-based tracer study system. It provides a detailed representation of the system's components, including the attributes and methods of each class, as well as the associations, dependencies, and interactions between them. Each class encapsulates specific responsibilities—such as handling user input, processing natural language queries, managing survey logic, or communicating with external services like the Google Sheets API.

By illustrating how these classes are interconnected, the diagram helps developers and stakeholders understand how data flows through the system, how objects collaborate to perform tasks, and how responsibilities are distributed across different parts of the architecture. This structural blueprint not only supports efficient system design and development but also facilitates future maintenance, scalability, and integration with additional services or features.

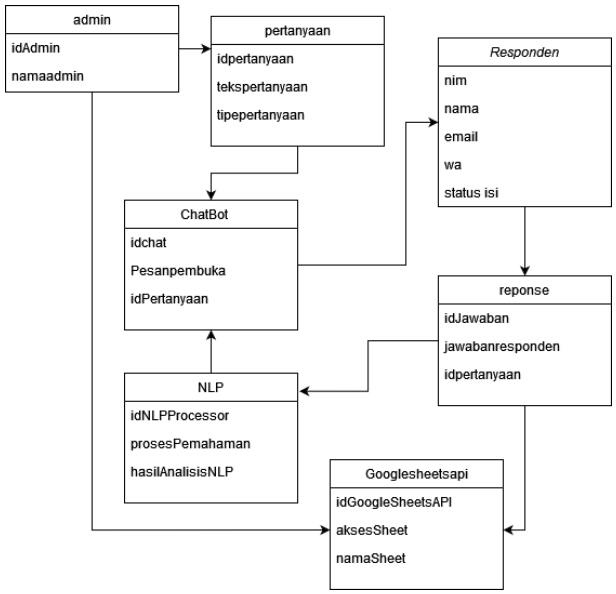

Figure 2. Class Diagram survey with whatsapp bot

This class diagram illustrates a chatbot-based tracer study system involving several main classes. The admin is tasked with managing survey questions stored in the Question class, which contains the identity, text, and question type. Respondents will receive questions through the ChatBot which acts as an intermediary to send opening messages and questions. Respondents fill out the survey, and their answers are stored in the **Response class,** which contains the identity of the

answer and its text. These answers are processed by NLP (Natural Language Processing), which analyzes the text for further understanding. Once processed, the answers are saved to the Google Sheets API, which automatically manages the saving of the survey data to Google Sheets for

documentation and further analysis. This diagram shows the communication and data processing flow, from the admin setting the questions to saving the respondents' survey results.

3.3. Sequence Diagram

This diagram shows the logical flow of a WhatsApp chatbot-based survey integrated with Google Sheets API to store responses and NLP to process possible text inputs. The interaction starts from the chatbot sending a notification to storing the final results from the respondents, providing a simple yet effective experience in collecting survey data.

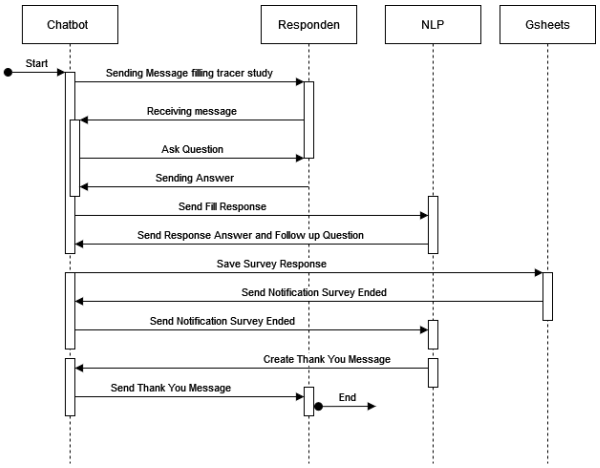

Figure 3. Sequence Diagram of survey with WhatsApp bot

The sequence diagram above illustrates the interaction process in a WhatsApp chatbot-based tracer study survey filling system. The process begins when the chatbot sends a message to the respondent to start filling out the survey. After receiving the message, the respondent answers the questions asked by the chatbot. Each respondent's answer is then forwarded and stored in the Google Sheets API. If necessary, the Natural Language Processing (NLP) system also processes the input from respondents to analyze the text sent. After all questions have been answered and the answers are saved, the chatbot sends a notification that the survey has been completed, and in closing, the chatbot sends a thank you message to the respondent. This diagram explains how the communication flow between the user, chatbot, Google Sheets API, and NLP takes place effectively to collect survey data with full automation.

3.4. Activity Diagram

This activity diagram illustrates the workflow in the process of filling out a tracer study using a WhatsApp chatbot integrated with Natural Language Processing (NLP) and Google Sheets API. The diagram shows how the chatbot interacts with respondents to collect survey answers, how NLP processes open-ended answers, and how survey results are stored in Google Sheets.

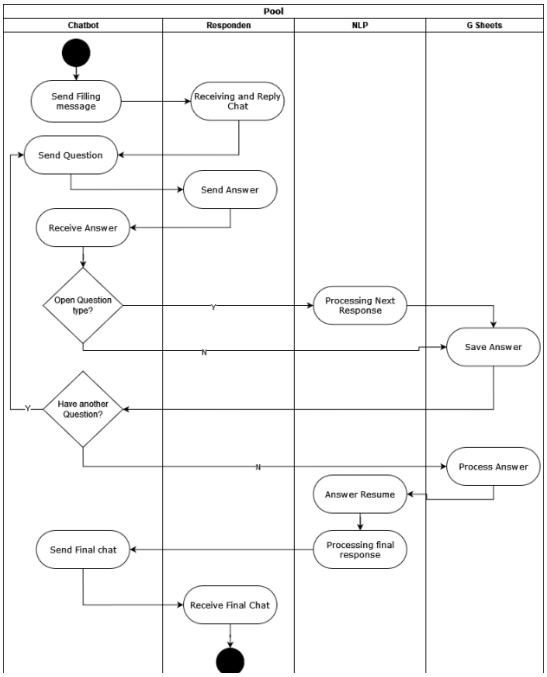

Figure 4. Activity Diagram survey with WhatsApp bot

This activity diagram illustrates the flow of filling out a tracer study via a WhatsApp chatbot, where the chatbot sends a filling message to the respondent. The respondent answers the questions, and if the answer requires further analysis, NLP processes it. The answers received, both from the respondent and the NLP results, are stored in Google Sheets using the API. The process continues until all questions have been answered, and ends with a thank you message from the chatbot. This diagram shows the interaction flow between the chatbot, respondents, NLP, and Google Sheets in collecting survey data.

4. Conclusion

In the discussion of this project, it can be concluded that the proposed system effectively addresses the challenges associated with collecting tracer study data by introducing an automated solution based on a WhatsApp chatbot. This chatbot is enhanced with Natural Language Processing (NLP) capabilities and is seamlessly integrated with the Google Sheets API to enable efficient data management. Technically, the chatbot serves as a communication interface that facilitates the bidirectional exchange of survey messages with respondents. NLP plays a critical role in automatically processing open-ended responses, enabling the system to interpret and categorize textual input without the need for manual intervention.

All responses—both structured and unstructured—are transmitted and stored directly into Google Sheets via API integration. This not only streamlines the data collection process but also supports easier data analysis and reporting. The entire system is built using a microservices architecture, which ensures that core components—such as the WhatsApp API, NLP service, and Google Sheets API—operate independently. This modular approach enhances the system's flexibility, scalability, and maintainability, making it easier to update or expand individual services without disrupting the entire application.

The flow of chatbot communication includes sending survey prompts to respondents, interpreting responses in real-time through NLP, and recording the final outputs into Google Sheets—all executed automatically without the need for human oversight. As a result, the system simplifies the operational process of tracer studies, significantly accelerates data collection, and minimizes human error during data processing.

Ultimately, the application of this integrated technology contributes to improving both the quality and quantity of tracer study data. This leads to more accurate evaluations of institutional performance and provides a solid foundation for designing and enhancing future educational programs.

References

[1] H. Schomburg, *Handbook for Graduate Tracer Studies*, 2nd ed. Moenchebergstrasse, Germany: International Centre for Higher Education Research, 2003.

[2] Kementerian Pendidikan dan Kebudayaan, "Tracer Study Kemdikbud," https://tracerstudy.kemdikbud.go.id/. Accessed: Oct. 15, 2024. [Online]. Available: https://tracerstudy.kemdikbud.go.id/

[3] Trisnani, "Pemanfaatan WhatsApp Sebagai Media Komunikasi Dan Kepuasan Dalam Penyampaian Pesan Di Kalangan Tokoh Masyarakat," *J. Komunika J. Komunikasi, Media dan Inform.*, vol. 6, no. 3, Dec. 2017, doi: 10.31504/komunika.v6i3.1227.

[4] Z. N. Baiti and F. Nugroho, "Aplikasi Chatbot 'Mi3' Untuk Informasi Jurusan Teknik Informatika Berbasis Sistem Pakar Menggunakan Metode Forward Chaining," *MATICS*, Oct. 2013, doi: 10.18860/mat.v0i0.2425.

[5] B. Setiaji and F. W. Wibowo, "Chatbot Using a Knowledge in Database: Human-to-Machine Conversation Modeling," in *Proceedings - International Conference on Intelligent Systems, Modelling and Simulation, ISMS*, IEEE, Jan. 2016, pp. 72–77. doi: 10.1109/ISMS.2016.53.

[6] J. Pustejovsky and A. Stubbs, *Natural language annotation for machine learning*. Sebastopol,USA: O'RellyMedia,Inc, 2013. [Online]. Available: http://it-ebooks.info/book/681/%5Cnpapers3://publication/uuid/906A922E-DE39-4CAB-8067-F222D065ACEF

[7] C. S. , Yulrio Brianorman, Rio Alamanda, "Aplikasi Pendeteksi Plagiat Terhadap Karya Tulis Berbasis Web Menggunakan Natural Language Processing Dan Algoritma Knuth-Morris-Pratt," *Coding J. Komput. dan Apl.*, vol. 4, no. 1, Jan. 2016, doi: 10.26418/coding.v4i1.13332.

[8] M. F. A. Muri, H. S. Utomo, and R. Sayyidati, "Search Engine Get Application Programming Interface," *J. Sains dan Inform.*, vol. 5, no. 2, pp. 88–97, Dec. 2019, doi: 10.34128/jsi.v5i2.175.

[9] Y. Y. Marwanta and B. Badiyanto, "mplementasi arsitektur microservice pada pembuatan surat unit kegiatan mahasiswa informatika dan komputer menggunakan Node.js," *JIKO (Jurnal Inform. dan Komputer)*, vol. 4, no. 2, pp. 71–80, Sep. 2019, [Online]. Available: https://ejournal.akakom.ac.id/index.php/jiko/article/view/516

[10] M. R. Adani, "Mengenal Metode Agile dan 12 Prinsipnya bagi Software," https://www.sekawanmedia.co.id/blog/metode-agile-development/. Accessed: Oct. 16, 2024. [Online]. Available: https://www.sekawanmedia.co.id/blog/metode-agile-development/

[11] I. Carolina and A. Rusman, "Penerapan Extreme Programming Pada Sistem Informasi Penjualan Pakaian Berbasis Web (Studi Kasus Toko ST Jaya)," *INOVTEK Polbeng - Seri Inform.*, vol. 4, no. 2, p. 157, Nov. 2019, doi: 10.35314/isi.v4i2.1043.