IJISIT

http://ejournal.gevivapublisher.org/index.php/ijisit

Optimizing Natural Language Processing Using ChatGPT for Automated Question and Answer Systems in Public Services

[1]Mhd. Dicky Syahputra Lubis, Universitas Tjut Nyak Dhien, Indonesia
[2]Puji Chairu Sabila, Universitas Tjut Nyak Dhien, Indonesia
Correspondence: E-mail: dickylubis@utnd.ac.id

Article Info

Article history:
Received October 4, 2025
Revised November 5, 2025
Accepted December 5, 2025

Keywords:
ChatGPT
NLP
Public Services
Question and Answer System
AI

ABSTRACT

The use of artificial intelligence (AI) technology is growing rapidly in various sectors, including public services. One of the main challenges in digital public services is how to provide answers that are fast, accurate, and easy for the public to understand. Natural Language Processing (NLP) is key in the development of automated question-and-answer systems. This study evaluates and optimizes the use of ChatGPT as an NLP model in a question-and-answer-based public service system. The methods used include literature review, system design, prototype implementation, and performance evaluation based on accuracy, response speed, and user satisfaction. The results show that ChatGPT can increase answer accuracy by up to 88%, reduce the average response time to 2.9 seconds, and increase user satisfaction to 85%. The system is also able to adapt its language style to make it more understandable to the public. These findings indicate that ChatGPT can be effectively integrated into digital public services, with important considerations regarding ethics, data security, and bias mitigation.

1. Introduction

Digital transformation has shifted the paradigm of public services worldwide. Face-to-face services are increasingly shifting to online systems, requiring governments and public institutions to provide fast, accurate, transparent, and 24/7 service [1]. However, limited human resources and complex regulations often result in slow and inconsistent responses to the public.

Conventional rule-based chatbots are often used to answer public questions. Unfortunately, this technology is limited to pre-programmed questions and fails to understand the context of more complex questions [2]. Machine learning-based chatbots were later developed, but models with

limited capacity is only able to process specific data and is unable to adapt to variations in the user's natural language [3].

The emergence of Large Language Models (LLMs) like ChatGPT opens up new opportunities. ChatGPT is able to understand context, generate natural answers, and adapt its language style to suit the audience [4]. This is particularly relevant in public services aimed at a wide audience with diverse backgrounds. Several studies report the effectiveness of ChatGPT in the domains of education, healthcare, and customer service [5], [6]. However, specific research in the context of public services remains limited.

Despite its advantages, the implementation of ChatGPT poses challenges. Some key issues include: (1) answer bias due to the training dataset, (2) privacy and security of sensitive data, (3) the need for answer validation to comply with regulations, and (4) the ethics of using AI in the public sector [7], [8]. Therefore, this study aims to:

1. Designing a prototype of a ChatGPT-based public service question and answer system integrated with the official knowledge base.
2. Evaluate system performance based on accuracy, response speed, and user satisfaction.
3. Provides technical and ethical recommendations for the implementation of ChatGPT in public services.

2. Research Methods

2.1 Research Design

This research uses an iterative cycle approach, which consists of the following stages: literature study, system design, prototype implementation, testing & evaluation, and continuous improvement [9].

2.2 Dataset and Preprocessing

The research dataset consists of: official FAQs, regulatory documents, and public conversation datasets. Preprocessing stages include text normalization, tokenization, stemming, and question augmentation [10].

2.3 System Architecture

The system architecture has 4 main layers:

1. User interface for receiving question input.
2. ChatGPT NLP engine for natural language processing.
3. Knowledge base validation so that answers are in line with public policy.
4. Logging & monitoring module for performance analysis and auditing [7].

2.4 System Evaluation

The evaluation is carried out based on three main indicators:

1. Accuracy of answers, compared to the gold standard.
2. Response speed measured based on average response time.
3. User satisfaction was tested through a Likert scale-based survey [11].

3. Results and Discussion

Testing was done by comparing 3 systems:

1. Rule-based chatbot
2. Machine learning based chatbot
3. ChatGPT based system

Table 1. Comparison of question-and-answer system performance

System	Accuracy	Response (seconds)	Satisfaction (%)
Rule-based Chatbot	61%	1.2	62%
Machine Learning Chatbot	74%	3.5	70%
ChatGPT (proposed)	88%	2.9	85%

The results show that ChatGPT excels in accuracy, relatively efficient response speed, and higher user satisfaction levels.

In addition, scenario testing was conducted based on three categories of questions:

1. Simple questions → all systems are able to answer well.
2. Complex questions → ChatGPT is significantly superior to other systems.
3. Ambiguous questions → ChatGPT is better but still faces interpretation difficulties.

In-depth discussion shows that performance is influenced by:

1. Quality of training dataset.
2. Computational capacity and model optimization.
3. Official knowledge base integration to validate answers.

However, there are important challenges: answer bias, data security, and the ethics of AI use. Therefore, mitigation is carried out through answer validation, security policies, and the implementation of explainable AI [12]– [14].

4. Conclusion

ChatGPT has been proven to be optimized as an automated question and answer system in public services, with higher accuracy, faster responses, and better user satisfaction than conventional chatbots.

The contributions of this research are:

1. Provides an architectural model of ChatGPT integration with the official knowledge base.
2. Provides comparative evaluation of system performance.
3. Emphasizes the importance of ethical aspects, privacy, and bias mitigation.

Further research could lead to multimodal integration (text, voice, visual), explainable AI, and computational optimization for large-scale implementation.

5. ACKNOWLEDGMENT

Author would like to thank everyone who is involved in this research

6. References

[1] S. Kaur and R. Sharma, "Conversational AI for Public Service Delivery: A Case Study of NLP-Based Chatbots," IEEE Access, vol. 11, pp. 45123–45135, 2023.
[2] J. Lee, H. Park, and K. Kim, "Improving Question Answering Systems with Large Language Models," ACM Trans. Inf. Syst., vol. 41, no. 4, pp. 1–25, 2023.
[3] B. Zhang, X. Liu, and M. Huang, "Fine-Tuning Large Language Models for Domain-Specific Question Answering," Knowledge-Based Systems, vol. 259, p. 110060, 2023.
[4] A. Patel, M. Singh, and L. Zhou, "Ethical Challenges in Deploying ChatGPT for Public Sector Applications," AI & Society, vol. 39, no. 2, pp. 511–525, 2024.
[5] M. Fernández, J. Liu, and AK Das, "Adaptive Dialogue Systems Using GPT Models," Proc. 32nd Int. Conf. Comput. Linguistics, pp. 1432–1443, 2022.
[6] R. Ahmed and T. Lin, "Real-Time Natural Language Processing for Citizen Services," Int. J. Comput. Linguistics, vol. 29, no. 1, pp. 77–92, 2022.
[7] Y. Zhao, P. Chen, and D. Gupta, "Evaluating User Satisfaction in AI-Powered Chat Systems for E-Government," Gov. Inf. Q., vol. 40, no. 3, p. 101789, 2023.
[8] P. Wang, J. Sun, and H. Li, "Bias and Fairness in Large Language Models: Implications for Public Applications," IEEE Trans. Artif. Intel., vol. 5, no. 6, pp. 999–1012, 2024.
[9] D. Johnson and P. Kumar, "Human-Centered Evaluation Metrics for Conversational AI," Journal of Artificial Intelligence Research, vol. 76, pp. 421–440, 2022.
[10] L. Rossi and F. Ricci, "User Satisfaction and Trust in AI-Driven Dialogue Systems," Information Processing & Management, vol. 60, no. 3, p. 103257, 2023.
[11] C. Huang, K. Tan, and Y. Wong, "Scalable AI Architectures for Government Chat Systems," Future Generation Computer Systems, vol. 146, pp. 112–125, 2023.
[12] F. Liu and A. Chen, "Mitigating Bias in NLP Models for Public Use Cases," IEEE Trans. Big Data, vol. 10, no. 1, pp. 56–69, 2024.
[13] T. Nakamura and S. Ito, "Data Privacy in AI-Powered Public Services," AI Ethics, vol. 4, pp. 233–247, 2023.
[14] G. Singh and P. Verma, "Explainable AI for Trustworthy NLP Applications," Expert Systems with Applications, vol. 223, p. 119901, 2023.