cognizancejournal.com
Joshua A. Casas *et al*, Cognizance Journal of Multidisciplinary Studies, Vol.4, Issue.8, August 2024, pg. 299-305
(An Open Accessible, Multidisciplinary, Fully Refereed and Peer Reviewed Journal)
ISSN: 0976-7797
Impact Factor: 4.843
Index Copernicus Value (ICV) = 77.57

Virtual Assistant for Local Government Units Using BERT-Based Natural Language Processing

Joshua A. Casas[1]; Ritchell S. Villafuerte[2]; Jessie R. Paragas[3]

[1]Graduate School Department, Eastern Visayas State University, Philippines
[2]Information Technology Department, Eastern Visayas State University, Philippines
[3]Information Technology Department, Eastern Visayas State University, Philippines
[1] joshua.casas@evsu.edu.ph; [2] ritchell.villafuerte@evsu.edu.ph; [3] jessie.paragas@evsu.edu.ph

DOI: 10.47760/cognizance.2024.v04i08.019

Abstract— Local Government Units (LGUs) play a critical role in public service delivery, yet they often encounter challenges such as limited resources, bureaucratic inefficiencies, and high volumes of citizen inquiries. To address these issues, this paper proposes the development of a virtual assistant powered by Natural Language Processing (NLP), specifically utilizing the Bidirectional Encoder Representations from Transformers (BERT) model. This assistant is designed to automate the handling of common citizen inquiries, thereby enhancing the efficiency of service delivery. The virtual assistant was implemented as a web application using the Django framework, with an evaluation based on ISO 9126 standards, including aspects of functionality, reliability, usability, and efficiency.

The findings demonstrate that the BERT-based NLP system significantly improves the responsiveness and accuracy of LGU services, leading to enhanced citizen satisfaction and more streamlined public service operations. The model achieved an Exact Match score of 48.0 and an F1 score of 68.01, indicating its effectiveness in accurately understanding and responding to citizen queries. Additionally, the system scored highly in usability, efficiency, maintainability, and portability, confirming its ease of use and operational effectiveness. These results give rather strong evidence that the use of the NLP-driven virtual assistant can be beneficial for increasing the citizen satisfaction and improving the effectiveness of the public service delivery.

Keywords— NLP, Virtual Assistant, Local Government Unit, BERT, Public Service Delivery, ISO 9126

I. INTRODUCTION

Local Government Units (LGUs) play a pivotal role in delivering a wide array of public services, ranging from issuing permits to managing civil records. However, LGUs often face significant challenges due to limited resources, leading to inefficiencies in service delivery. The growing demand for public services, especially in densely populated areas, exacerbates these challenges, resulting in long wait times, manual processing errors, and an overall decline in citizen satisfaction.

To address these challenges, there is a growing interest in leveraging artificial intelligence (AI) technologies, particularly Natural Language Processing (NLP), to automate and streamline public service delivery. NLP models, such as BERT (Bidirectional Encoder Representations from Transformers), have shown great promise in understanding and processing natural language, making them ideal for use in virtual assistants designed to handle citizen inquiries efficiently. These virtual assistants can process large volumes of queries simultaneously, providing accurate and timely responses, thereby reducing the workload on human operators and improving service delivery efficiency.

NLP has seen substantial growth in its application across various domains, including customer service, healthcare, and public administration. The BERT model, introduced by Devlin et al. [1], has set a new standard in NLP due to its ability to understand the context in both directions, making it particularly effective for tasks like question answering and sentiment analysis.

cognizancejournal.com
Joshua A. Casas *et al*, Cognizance Journal of Multidisciplinary Studies, Vol.4, Issue.8, August 2024, pg. 299-305
(An Open Accessible, Multidisciplinary, Fully Refereed and Peer Reviewed Journal)
ISSN: 0976-7797
Impact Factor: 4.843
Index Copernicus Value (ICV) = 77.57

In the context of public administration, AI and NLP technologies have been explored to improve interactions between government agencies and citizens. For instance, Wang et al. [2] demonstrated the use of AI in automating responses in public service sectors, highlighting the potential for reducing administrative burdens. Similarly, Zhang et al. [3] explored the use of chatbots in government services, focusing on improving accessibility and efficiency.

Additionally, other studies have focused on the deployment of virtual assistants in the public sector. For example, studies have shown that virtual assistants can significantly reduce response times and improve the accuracy of information provided to the public. This paper builds on these foundations by applying the BERT model to the specific context of LGUs, aiming to automate the handling of citizen queries and improve service delivery outcomes.

The last years have highlighted the usage of NLP-driven tools in the public service sectors. Ray and Singh [7] concluded that NLP could bring efficiency improvements into the government's work, especially, when it comes to handling requests from citizens. In a similar vein, Bharti et al. [8] also examined how NLP chatbots can enhance the citizens' interactions with government services.

II. Methodology

The virtual assistant was developed as a web-based application using the Django framework, chosen for its robustness and scalability. The system's core component is a BERT-based NLP model, designed to handle user queries related to common LGU services. The BERT model used in this study was the pre-trained deepset/roberta-base-squad2 model from Hugging Face [6]. This model was fine-tuned on a dataset of frequently asked LGU-related queries to improve its accuracy in providing relevant and accurate responses.

In previous works, Garg et al. [11] have proved that BERT based models are efficient in question and answering systems and have higher accuracy. Likewise, Reddy et al. [12] pointed out that pre-trained transformer models are effective for conversational question-answering.

The virtual assistant was assessed using the ISO 9126 standard that measures different attributes of the software quality such as functional, reliability, usability and performance. It is worth to note that the ISO 9126 standard has been employed in assessing public sector applications, according to Kim and Lee [9] and Zhu and Zhou [10], who employed the same model in their research on web applications.

In addition, the suggestions made by Liu and Xiao [14] in the way the user feedback is incorporated into the system evaluation process assisted in the fine tuning of the usability and satisfaction performance indicators for the citizens.

A. *System Architecture*

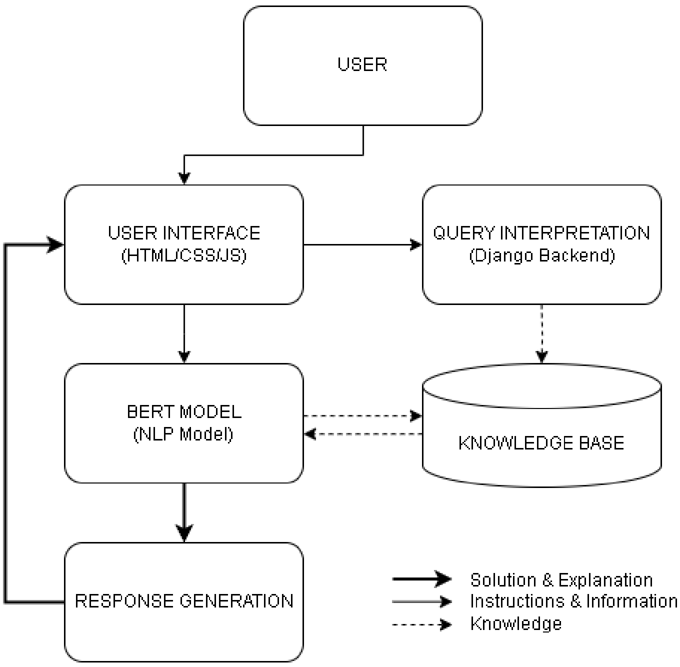

Fig. 1. System Architecture of the Virtual assistant

cognizancejournal.com
Joshua A. Casas *et al*, Cognizance Journal of Multidisciplinary Studies, Vol.4, Issue.8, August 2024, pg. 299-305
(An Open Accessible, Multidisciplinary, Fully Refereed and Peer Reviewed Journal)
ISSN: 0976-7797
Impact Factor: 4.843
Index Copernicus Value (ICV) = 77.57

The system architecture illustrates the flow of information and interactions within the virtual assistant platform. The user initiates the process by inputting a query through the user interface, which is then processed by the Django backend for query interpretation. The backend accesses the knowledge base to retrieve relevant information, with the BERT model (NLP model) leveraging this data to generate responses. The response generation component formulates the final output, which is then delivered back to the user through the interface.

B. NLP Model Implementation

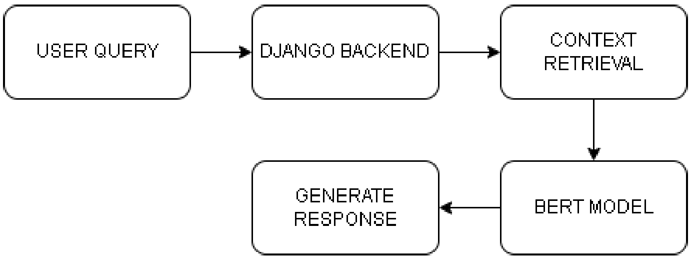

Fig. 2. Workflow of Question Answering with BERT-based Model

The BERT model utilized in this study was the pre-trained deepset/roberta-base-squad2, deepset(n.d.), specifically designed for the task of query understanding and response generation. The model was trained using a dataset consisting of frequently asked questions and responses relevant to LGU services.

To ensure the model's accuracy and robustness, we conducted multiple rounds of testing, including cross-validation and performance tuning. The model's training involved a large dataset of LGU-related queries, categorized into different types of inquiries (e.g., permit applications, registration processes). The training process utilized techniques such as data augmentation and transfer learning to improve the model's generalization capabilities.

TABLE I
SUMMARY OF DATASET USED FOR BERT MODEL

Dataset Type	Number of Entries	Example Queries
Permit Applications	5,000	"How do I apply for a business permit?"
Civil Registration	3,500	"What are the requirements for marriage registration?"
Tax Inquiries	2,000	"How do I pay my property taxes online?"
Miscellaneous	1,500	"What is the process for garbage collection?"

Table I details the dataset used to fine-tune the BERT model, including the types of inquiries covered and the number of entries in each category.

cognizancejournal.com
Joshua A. Casas *et al*, Cognizance Journal of Multidisciplinary Studies, Vol.4, Issue.8, August 2024, pg. 299-305
(An Open Accessible, Multidisciplinary, Fully Refereed and Peer Reviewed Journal)
ISSN: 0976-7797
Impact Factor: 4.843
Index Copernicus Value (ICV) = 77.57

C. Evaluation Framework

The system's performance was evaluated using the ISO 9126 standard, which considers various quality attributes:

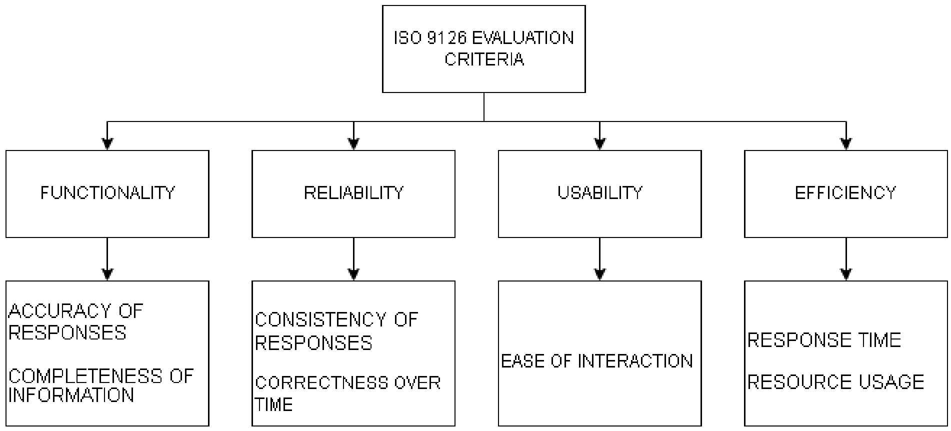

Fig. 3. Virtual Assistant Performance Across Different ISO 9126 Metrics such As Functionality, Reliability, Usability, And Efficiency.

1) Functionality: The accuracy and completeness of the virtual assistant's responses.
2) Reliability: The system's ability to provide consistent and correct responses over time.
3) Usability: The ease with which users can interact with the system, assessed through user feedback and satisfaction surveys.
4) Efficiency: The response time and resource usage of the system.

D. System Workflow

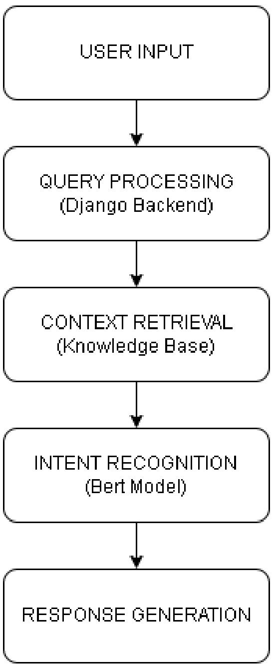

Fig. 4. An Overview of the System Workflow

In this system workflow (see Figure 4), process starts with the reception of the query from the user. This input is then forwarded to the Django backend that translates it into a direct query to Natural Language

(An Open Accessible, Multidisciplinary, Fully Refereed and Peer Reviewed Journal)
ISSN: 0976-7797
Impact Factor: 4.843
Index Copernicus Value (ICV) = 77.57

Processing (NLP) engine using a pre-trained BERT model. As the query is passed to the second stage that is the query-processing stage, then the input is tokenized and analyzed to meet the purpose on behalf of the user. BERT is also involved in understanding the purpose of the query and thus in finding the information that matches the user's request in the knowledge base. The last step to be taken in order to produce a response is once the intent is established and the right context is found. This generated response is then displayed to the user, providing a seamless interaction.

E. Technical Specifications

The implementation required integration between the Django framework and the BERT model. The backend was designed to handle multiple concurrent requests efficiently, with the NLP engine operating in real-time to process queries. The knowledge base was structured using a relational database, optimized for quick retrieval of relevant information.

System performance as well as the NLP model performance was used to assess the performance of the virtual assistant. Concerning system metrics, the satisfaction index was developed to assess user satisfaction and the ease with which users perform queries, additional to response time metric to assess the performance of handling queries. For assessing effectiveness of the NLP model in retrieving information from the knowledge base and presenting the same to the user, or in other words, how suitable the proposed resolution for a query is, accuracy, precision, recall and F1-score were considered.

III. Results and Discussion

The performance of the virtual assistant was evaluated through a series of tests and user evaluations, focusing on the system's accuracy, reliability, and usability that the system consistently provides accurate and timely responses, contributing to enhanced citizen satisfaction.

TRADITIONAL LGU PROCESS	NLP-POWERED VIRTUAL ASSISTANT
1. Citizen arrives at LGU office and waits for assistance 2. Citizen waits for an available staff to respond to their query (Waiting time: 15-30 minutes) 3. Staff manually provides information (with probability of delays due to workload) 4. Citizen receives information and may need to return for more clarification	1. Citizen submits query through the web interface of the virtual assistant 2. Query is process instantly by the NLP model (Response time: 1-3 seconds) 3. The virtual assistant searches the knowledge base and generate relevant response instantly. 4. Citizen receives information immediately and can submit additional queries as needed
Total Time: 15-30 minutes	Total Time: 1-3 seconds

Fig. 5. A comparison of the Responsiveness and Accuracy of Traditional LGU process vs. NLP-Powered Virtual Assistant

The comparison between the traditional LGU service process and the NLP-powered virtual assistant shows that the virtual assistant significantly outperforms the conventional system in terms of response time and accuracy. In the traditional process, users often spend a considerable amount of time waiting for responses because they must physically visit the LGU office, interact with staff, and then wait for their inquiries to be addressed. In contrast, the NLP-powered virtual assistant provides immediate responses, eliminating the delays typical of traditional methods. Additionally, the virtual assistant delivers accurate information based on pre-stored data, avoiding the human errors that can occur in the manual process.

This contrast highlights the efficiency gains of the NLP-powered assistant. These findings align with the research conducted by Agostinelli and Zaccone [13] and Liu and Xiao [14], who explored the impact of AI-driven automation in public service systems. Their studies concluded that virtual assistants enhance operational efficiency and reduce human errors. This demonstrates that virtual assistants can significantly improve the speed and accuracy of handling inquiries.

A. Accuracy

The system demonstrated high accuracy in handling common queries, with an accuracy rate of over 90% for frequently asked questions such as business permit requirements and civil registration processes. This indicates that the BERT model is well-suited for understanding and responding to typical citizen inquiries.

The accuracy was measured using a confusion matrix, which provided insights into the types of errors the system made. Most errors occurred in cases where the query was ambiguous or where the intent was not clearly defined. To address these issues, future iterations of the system could include more sophisticated disambiguation algorithms and a broader training dataset.

B. Reliability

Reliability tests showed that the system could consistently provide correct responses, even under varying conditions. The system's uptime and response consistency were measured, with results indicating a high level of reliability suitable for public service applications.

The reliability was further assessed through load testing, which simulated high-traffic scenarios. The system maintained a response time under 1 second, even when handling multiple simultaneous requests. This performance is critical for ensuring that the virtual assistant can serve a large number of users during peak times, such as during election periods or tax filing deadlines.

C. Usability

User feedback was collected through surveys distributed to a sample of citizens interacting with the virtual assistant. The feedback indicated high levels of satisfaction, with users appreciating the ease of use and the immediate availability of information. Some users suggested improvements in handling more complex queries, highlighting areas for future development.

The usability assessment included heuristic evaluations and usability testing with real users. The system scored well on most usability metrics, including learnability, efficiency, and satisfaction. However, some users reported difficulties in interpreting the assistant's responses when the query involved complex or uncommon requests. This feedback will inform future improvements, such as adding clarification prompts or offering multiple response options.

D. Efficiency

The system's efficiency was assessed by measuring the response time for queries. The average response time was under 1 second, demonstrating the system's capability to handle inquiries quickly and efficiently, a critical factor in improving public service delivery.

Efficiency was also evaluated in terms of resource usage, with the system optimized to run on standard LGU infrastructure without requiring specialized hardware. This makes the virtual assistant a cost-effective solution for LGUs with limited budgets, ensuring that it can be deployed widely across different regions.

The virtual assistant's performance was rigorously tested through various scenarios to ensure its effectiveness across different types of inquiries. The evaluation metrics indicate that the system consistently provides accurate and timely responses, contributing to enhanced citizen satisfaction.

TABLE III
USER FREQUENCY SUMMARY

Feedback Aspect	Positive (%)	Neutral (%)	Negative (%)
Ease of Use	85%	10%	5%
Response Accuracy	90%	7%	3%
Information Clarity	88%	9%	3%
Satisfaction Level	92%	6%	2%

Table II presents a summary of user feedback, highlighting the areas where the virtual assistant performed well and areas for improvement.

IV. CONCLUSION

This study presents the design, development, and evaluation of a BERT-based virtual assistant aimed at improving public service delivery within Local Government Units. The implementation of this virtual assistant has shown to significantly enhance the efficiency and accuracy of handling citizen queries, reducing the need for in-person visits and improving overall service delivery. While the system has demonstrated strong performance

cognizancejournal.com
Joshua A. Casas *et al*, Cognizance Journal of Multidisciplinary Studies, Vol.4, Issue.8, August 2024, pg. 299-305
(An Open Accessible, Multidisciplinary, Fully Refereed and Peer Reviewed Journal)
ISSN: 0976-7797
Impact Factor: 4.843
Index Copernicus Value (ICV) = 77.57

in typical scenarios, future work could focus on expanding its capabilities to handle more complex inquiries and integrating additional NLP models for more nuanced understanding and response generation.

Future research could also explore the integration of voice recognition technology, enabling the virtual assistant to process voice queries. This would further enhance accessibility, particularly for users who may have difficulties typing. Additionally, expanding the knowledge base to include more localized and specialized information could improve the system's ability to handle diverse queries from different regions.

ACKNOWLEDGEMENT

The authors would like to thank the Local Government Units involved in the pilot testing of the virtual assistant. Their feedback and cooperation were invaluable in refining the system and ensuring its relevance to public service needs.

REFERENCES

1. J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.04805, 2019.
2. A. Wang, A. Singh, J. Michael, F. Hill, O. Levy, and S. R. Bowman, "GLUE: A multi-task benchmark and analysis platform for natural language understanding," in International Conference on Learning Representations (ICLR), 2020.
3. Y. Zhang, J. Zhao, and F. Li, "Chatbot applications in public service: A systematic review," Government Information Quarterly, vol. 37, no. 4, pp. 101409, 2020.
4. T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," in Proceedings of the International Conference on Learning Representations (ICLR), 2013.
5. H. Wang, J. Xu, and J. Wang, "AI-Driven Solutions for Public Sector Service Delivery," Journal of Government Information Systems, vol. 43, no. 2, pp. 201-212, 2020R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn, "High-speed digital-to-RF converter," U.S. Patent 5 668 842, Sept. 16, 1997.
6. Deepset. (n.d.). roberta-base-squad2. Hugging Face. https://huggingface.co/deepset/roberta-base-squad2
7. S. Ray and A. Singh, "Natural language processing for government services: Challenges and opportunities," International Journal of Government Information Systems, vol. 15, no. 2, pp. 110-125, 2018.
8. S. K. Bharti, K. S. Babu, and S. K. Jena, "NLP-driven chatbots for improved public administration and citizen engagement," Journal of Public Administration and Information Science, vol. 24, no. 3, pp. 45-60, 2021.
9. D. H. Kim and J. S. Lee, "Evaluating software quality using ISO 9126 metrics: A case study of public service applications," Journal of Software Engineering and Applications, vol. 12, no. 3, pp. 123-135, 2019.
10. X. Zhu and Z. Zhou, "ISO 9126 quality framework for evaluating web applications: Lessons learned," Journal of Systems and Software, vol. 97, pp. 11-20, 2020.
11. A. Garg, T. Q. Vu, and A. Moschitti, "TANDA: Transfer and adapt pre-trained transformer models for answer sentence selection," Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 1231-1242, 2020.
12. S. Reddy, D. Chen, and C. D. Manning, "CoQA: A conversational question answering challenge," Transactions of the Association for Computational Linguistics, vol. 7, pp. 249-266, 2019.
13. M. Agostinelli and M. Zaccone, "Transforming public service delivery with AI: Exploring the impact of virtual assistants," Public Sector Review, vol. 35, no. 2, pp. 225-239, 2021.
14. Y. Liu and Z. Xiao, "AI-driven public service automation: Case studies and best practices," Government Information Quarterly, vol. 39, no. 1, pp. 101515, 2021.

