International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Transformative AI for Healthcare Customer Service: Auto-Documentation of Compliance-Critical Questionnaires

Leela Naga Sai Vamsi Krishna Dogiparthi

University of Bridgeport, USA

Abstract

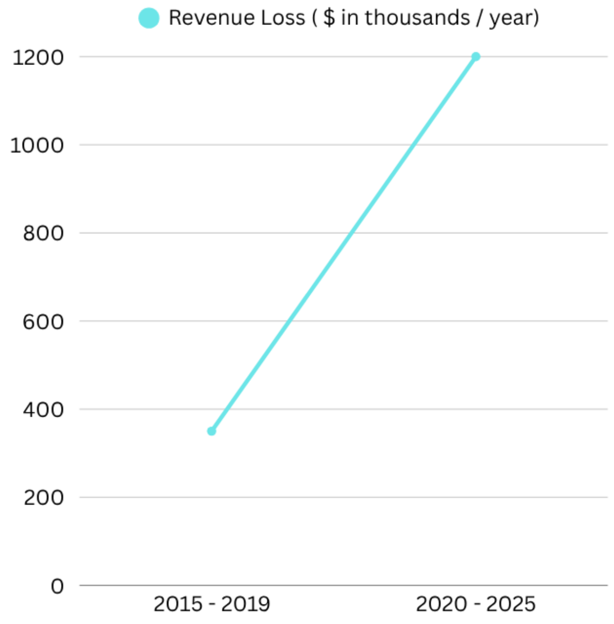
Over the last two decades, the population across the world has increased exponentially so as the number of healthcare requests. Especially, post-COVID the number of customer service calls for various healthcare related queries across the healthcare providers has increased by 65% [1]. The staffing shortages have made it harder for the providers to answer every call in a defined time. Also, 60% of users reported that they have higher waiting times, which is very frustrating and doesn't leave a pleasant impression on the users' mind. In our survey with customer care representatives across the various healthcare providers showcased that the majority, 75% - 80% of the time spent on a call is taken to document compliance critical questions which are manual and tedious. This paved path to our study on developing an AI-powered live transcription agent that simultaneously converts voice interactions to text, extracts answers to predefined clinical/billing questions, and provides real-time feedback to healthcare customer service agents regarding unanswered queries. This work establishes a new paradigm in the healthcare customer service realm with possibilities of transforming the healthcare infrastructure across the nation.

Keywords: Natural Language Processing (NLP), Natural Language Understanding (NLU), Sentiment Analysis, User Intent Recognition, Speech to Text, Prompt Engineering, Plugins, Responsible AI, Encryption at rest, Compliance, HIPAA.

1. Introduction

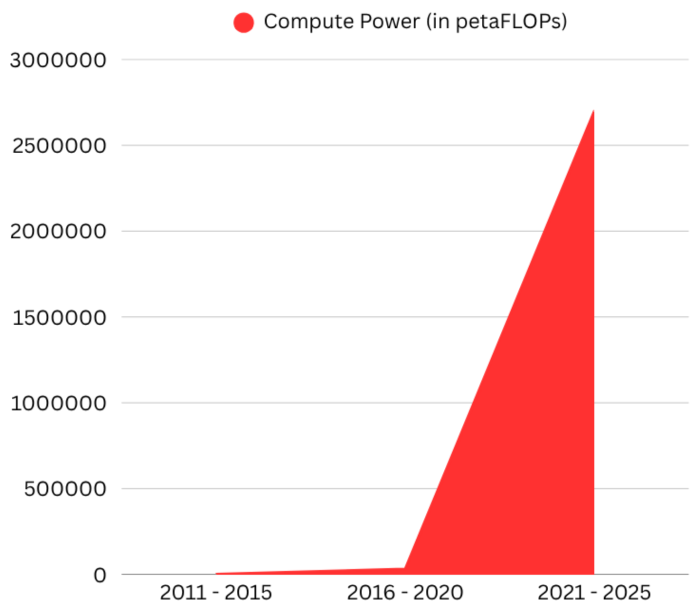
The call volume across healthcare providers has increased by 230% [2] over the past decade. Making it harder for already short-staffed providers to make all the calls within a satisfactory turnaround timeframe. According to the many surveys 65% of the patients [1] these days are expecting high standards from healthcare providers which only made things worse for the providers. According to our survey with customer service representatives across healthcare providers, it has been expressed that most of the time spent during a call with patients is spent documenting the compliance, system specific questionnaire and summarizing the call transcript during post call transcription. It's measured that 75% - 80% of call time is related to the tedious manual tasks mentioned above. These tedious tasks have very direct correlations with call wait times that patients observe these days. Over 40-45% [3], [4] of the patients has reported that they waited more than 30 minutes to talk to a healthcare customer representative. Even after implementation of callback systems, 25% of patients abandoned the calls after 15 minutes. [3], [4]. These calls have direct correlation to the satisfaction rates where >30% of the patients who waited for more than 15mins have reported lower satisfaction, even after being seen by a doctor. All these have resulted in revenue loss of ~1.2M/year over the period of 2020-2025 [3], [4].

Figure 1: Revenue loss in thousands of dollars



In the current digital & Artificial intelligence era resolving this issue and converting the manual process to an automated process with very high accuracy became very much possible. Especially when you can design a scalable and adaptable Artificial Intelligence agent and with the paradigm shift in the computing power of artificial intelligent LLM model over the decade, it is very possible to solve this problem with ease. Our study concentrates on the design and implementation of novel transformative agentic framework to automate the manual healthcare documentation process with real-time insights to the customer service representative to improve the quality of the patient care experience with higher positive call rates with lower wait times. Our agentic framework is expected to reduce the overall manual process by 70% and call wait times by 75%. Using pre-trained LLM models the organization is going to save 100% on training and maintenance costs.

Figure 2: Computing power over past decade [5], [6]



2. Related Work

Over the past decade there were many studies conducted in sentiment analysis, speech recognition, speech to text conversion and several methodologies have been proposed. Below are some of the prominent studies.

Article 1 authors [7] conducted their study on performing sentiment analysis on online product reviews using machine learning techniques such as Random Forest, Logistic Regression, k-Nearest Neighbor, and Catboost classifier to improve the outcome of sentiment analysis. This paper also emphasizes some of the data pre-processing, feature extraction and polarity techniques. This paper mainly concentrates on using custom trained models for sentiment analysis which requires training, publishing & maintenance costs. The techniques proposed by authors do have a steep learning curve. The study doesn't concentrate on the usage of some pre-trained LLM models to replace custom trained models for better scalability, performance and equal or higher accuracy with lower costs.

Article 2 [8] authors conducted their study on performing sentiment analysis on product reviews in E-commerce domain. The paper emphasizes the need for advanced machine learning models such as dynamic Brand-Topic Models (dBTM) for tracking sentiment on a specific brand overtime. Aspect-Based Sentiment Analysis (ABSA) to associate sentiment with specific aspects of the brand. The study also proposes Sentiment Mapping strategy to identify the strengths and weaknesses of a brand to provide quality product insights. Their proposed system also utilizes Continuous Learning with Naïve Bayes for enhanced classification accuracy. Even with extended study the authors only concentrated on specific E-commerce business cases, which makes their model and proposed methodologies less adaptable to other domains. Along with it, their study doesn't concentrate on how Pre-trained LLM models can replace all the custom models and the costs associated with it.

Article 3 [9] authors conducted their study on similar sentiment analysis on product review in the E-Commerce sector. Their study emphasized a fundamental problem of sentiment analysis i.e. sentiment polarity categorization. Their proposed system tackles a specified problem in phases which includes Sentiment Phrase extraction with an algorithm for negation phrases identification, Sentiment score computation, feature vector generation and sentiment polarity categorization. Their study is primarily and deeply focused on solving the polarity problem of various product reviews, which makes their study less adaptable to other industries. Also, the paper doesn't emphasize leveraging the pre-trained LLM models with Prompt engineering & RAG pattern to produce a system which is more adaptable and reusable.

Article 4 [10] authors performed their research on using a new version of Long Short-term memory RNN called deep LSTM RNN which improves the speech to text accuracy even with the fact of speech recognition being dynamic in nature. Their study focused on implementing deep neural networks of bi-directional RNN with continuous feedback mechanisms. Their neural networks implement RNN transducer, decoding and regularization to improve the performance of the neural networks by reducing the amount of information transmission. Even though being a novel approach, their proposed model has shown significant improvement in speech recognition but require extensive custom training and maintenance. Which might not be feasible for every adopting organization or industry.

Article 5 [11] showcases the important study on extracting the answers to the questions from the information provided to the model. The authors explained various important stages involved in identifying a relatively accurate answer to the question from the information provided as a document or text. They have emphasized TF-IDF similarity, Jaccard index, Word decorations, text deep learning similarity algorithms for mapping a question to a possible answer. Their proposed system identifies the various ways in which the query formulation can result in better outcomes especially with entity extraction. They also showcased their study on different types of question answering models and how they work in finding answer to the presented question. Their study doesn't reflect on utilizing the powerful capabilities of some of the pre-trained LLM models for extracting answers with lower maintenance & training costs. Which can result in higher accuracy with little overhead for the adopting systems.

Article 6 [12] proposed an Open-source speech to text software which utilizes some of the widely used modules such as Voice Activity Detection (VAD) and automatic speech recognition (ASR) with 100% accuracy and lower Word error rate. Being an open-source system, their software has higher adoptability rates, but their proposed architecture and software mainly concentrates on digital forensics investigation which limits its adaptability into different domains.

3. Methodologies

Our automated agent framework implements below a set of methodologies for achieving the end goal of the agent. Each methodology plays a vital role in the overall agent framework implementation.

1. Sentiment Analysis.
2. Summarization.
3. Natural Language Processing (NLP).
4. Natural Language Understanding (NLU).
5. User Intent Recognition.
6. Speech to Text.
7. Retrieval Augmented Generation (RAG).
8. Q&A Extraction.
9. Prompt Engineering.
10. Plugins/Tools.
11. Encryption at rest.
12. Responsible AI Practices.
13. JSON Serialization & Deserialization.
14. Remote Procedure Calls (RPC).

Sentiment Analysis: The agent performs live sentiment analysis on the overall conversation between the customer and customer service agent [13]. The scores are calculated for three categories positive, negative and neutral. At the end of the call the sentiment analysis scores of each category are finalized and the dominated feeling becomes the overall sentiment of the conversation. This feature allows the agent and healthcare organization to perform quantifiable measurements on total positivity across the cases served by each agent to understand and decide on the next steps to improve the system. The system implements Continuous Learning with Naïve Bayes (CLNB) via pre-trained LLM models [14] for improved accuracy.

Summarization: The call transcript (text) produced from audio for the current conversation is summarized at the end of the call and presented to the customer care representative to make final edits to the summary and submit. The summarization includes Entity extraction, key points of the conversation, speaker recognition, next set of action items. This feature allows the representative to quickly submit the summary and reduce the time spent on manually noting the conversation points and summarizing it at the end. This feature is expected to drastically reduce the post-call work done by each representative and decrease the waiting time for new callers.

Natural Language Processing (NLP) & Natural Language Understanding (NLU): Natural Language Processing enables us to process the text, understand and extract the entities out of call text, understand the semantic meaning of the sentences and words. These methodologies play a pivotal role in the overall functioning of the agent. NLP & NLU is used to understand the user intent for injecting intent specific questionnaires into the LLM model for extraction and further processing.

User Intent Recognition: Based on the live audio transcript the LLM model identifies the user intent to fine tune the model behavior and to identify the set of questions to be injected into the LLM model. This user intent recognition plays a vital role in the overall implementation of RAG pattern with LLM system prompts.

Speech to Text: The live audio is transcribed into text as the call progress. This allows the customer service agent to quickly verify the conversation happening between them. The speech to text [15] is fed into the LLM model to perform further sentiment analysis, entity extraction, question and answers, sentiment scoring, speaker recognition, key point extraction and call summarization. [10], [12]

Q&A Extraction: LLM models are trained to read specific sets of questions that must be answered by every caller for further processing. They are trained to process the live call transcript for auto extracting answers to the matching questions.

Retrieval-Augmented Generation (RAG): RAG is a very powerful feature, which allows us to inject system specific real-time information into pre-trained LLM models to not hallucinate and provide grounded responses. In our agent, we are using RAG to inject the system specific questionnaire and example expected answers to the LLM model. RAG pattern was also used to inject appointment information, location information, system specific verbiage etc.

Prompt Engineering: We have implemented 3S principles in designing our prompts that define the behavior the agents in the system. Using prompt engineering, you decide how the LLM should behave and to help it understand the needs of the system. Using this technique, we eliminate the use of custom trained models for each use case but use more generalized models for most of the system operations. This technique saves thousands of dollars that organizations must spend to train, develop, deploy and maintain custom models.

Plugins/Tools: LLM models are equipped with powerful features of calling certain plugins/tools during their overall operation flow. Using plugins, we can easily inject system specific data that is readily available in the consumer system, thus enable the ability to implement RAG pattern. Plugins are also used as the pre and post processing filters for safeguarding the underlying model with harmful and inappropriate prompts/data.

Encryption at rest: The system implements encryption at rest protocol, to encrypt the audio/text transcripts to safeguard the caller private information and comply with HIPAA and other privacy standards.

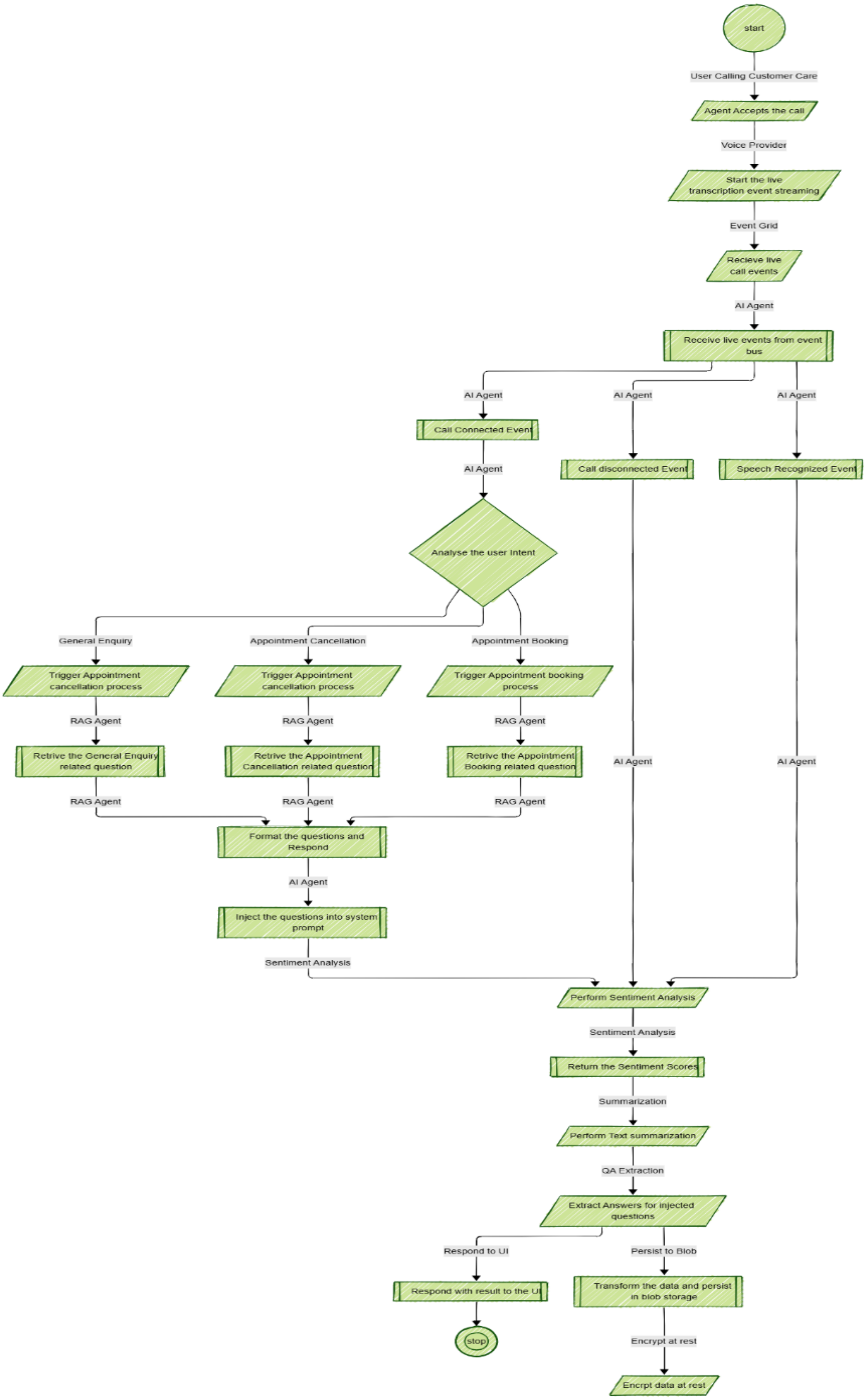
Responsible AI Practices: The system is developed using ethical/responsible AI practices such as inclusiveness, accountability, security, privacy etc. to emphasize call to action on human centered AI agent design system to enable reliability and fairness. Using responsible AI practices is a must and should be incorporated into every phase of the AI agent development process.

JSON Serialization & Deserialization: The system implements JSON serialization and deserialization techniques for converting the audio/text transcription for easier injection and transportation between the different layers of the internal systems. JSON is one of the industry standard protocols for communication between systems both internal & external. Implementation of JSON improves the scalability and adaptability of the agent.

Remote Procedure Calls (RPC): Remote Procedure Calls (RPC) enables the system to invoke and communicate data between client and server. This allows client and server layers to interact and execute procedures asynchronously on demand. Using RPCs the system pushes the live audio transcript to the customer care representative with extracted answers to required questions, and an end call summary with key points, next actions etc.

HIPAA: The system has been designed to comply with HIPAA to safeguard the user's privacy and personal information.

Figure 3: AI Agent Flow Diagram

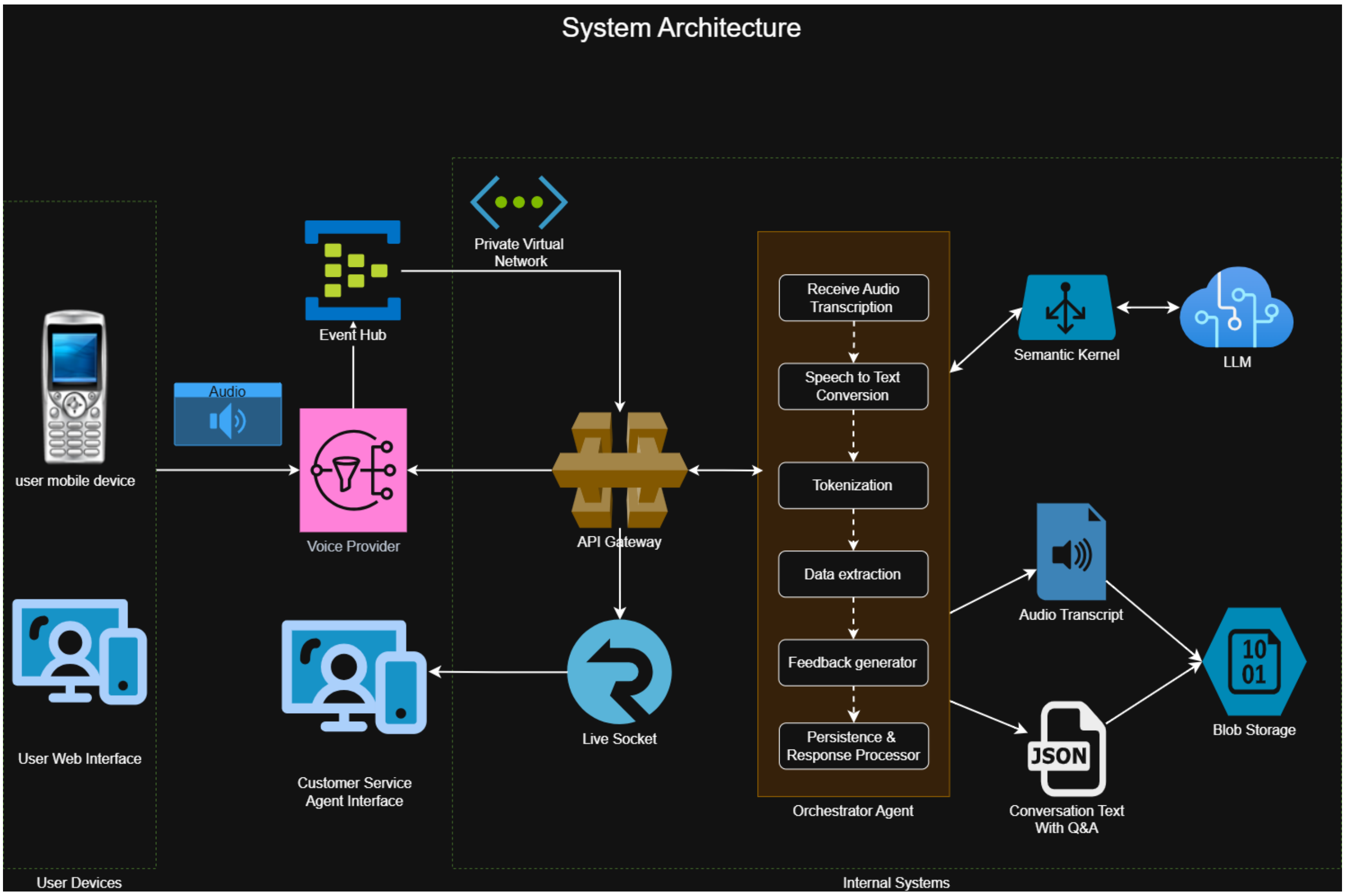


4. Proposed System

The proposed AI agent system implements speech to text methodology [12], [15] to convert the live audio feed to text for further processing. System implements continuous sentiment analysis using Continuous Learning with Naïve Bayes (CLNB) for improved accuracy on the converted to text to score the overall user sentiment point in time of the conversation. At the end of the conversation a final set of scores is calculated and overall dominated sentiment. System also natural language processing (NLP) & natural language understanding (NLU) to understand and extract the key information from the audio transcript. These techniques enable the system to perform entity extraction, question & answer pair mapping, tokenization. Natural language generation (NLG), deep neural network is used to produce feedback responses for follow up prompts to Customer care representatives to help them improve quality of the call. Each layer of the system is implemented using Responsible AI practices to ensure that the system follows ethical standards of AI development. The proposed system implements plugins and filters to implement RAG pattern for injecting system specific information for producing grounded responses as well as implement security on the system injected prompts. All the system audio and text transcripts are encrypted at rest and persisted inside the blob storage.

5. System Architecture

Figure 4: Automation Agent System Architecture



6. Results:

Sentiment Analysis: Our agent system showcased about ~94% accuracy in identifying the caller's overall sentiment. The confusion matrix shows the end results.

Figure 5: Sentiment analysis results

Sentiment Results				
	Positive	Negative	Neutral	SUM
Positive	98 31.21%	0 0.00%	2 0.64%	100 98.00% 2.00%
Negative	5 1.59%	94 29.94%	6 1.91%	105 89.52% 10.48%
Neutral	1 0.32%	4 1.27%	104 33.12%	109 95.41% 4.59%
SUM	104 94.23% 5.77%	98 95.92% 4.08%	112 92.86% 7.14%	296 / 314 94.27% 5.73%

Figure 6: Sentiment analysis recognition prompt template

````
name: ConversationSentimentDeterminationTemplate
template: |
  <message role="system">
   # System Instructions:
   You are intelligent language knowledge expert, who can analyze customer sentiment based on the transcribed conversation between Customer and Representative
   You sentiment analysis response should be in json format.
   The response should include Customer Conversation summary, Is overall conversation is positive or negative or neutral, Sentiment_Postive score, Sentiment_Negative score, sentiment_neutral score.
   # Ouput Instructions:
   # Example Response: Only provide the response in json format. No markdown or any other format. No comments.  Do not include any other information or comments in the response.
   {
     "sentiment_positive_score": 0.9,
     "sentiment_negative_score": 0.05,
     "sentiment_neutral_score": 0.05,
     "overall_sentiment": "positive",
   }
   # Example response 2:
    {
      "sentiment_positive_score": 0.1,
      "sentiment_negative_score": 0.8,
      "sentiment_neutral_score": 0.1,
      "overall_sentiment": "negative",
    }
   - No markdown or comments shoulbe in the response. no ```json should be included.
   - Only respond in pure json.
  </message>
  <message>
   {{CustomerConversation}}
  </message>
template_format: handlebars
description: This template is to perform sentiment analysis on Customer Conversation and provide a response based on the sentiment.
input_variables:
  - name: CustomerConversation
    description: Customer Conversation.
    is_required: true
    allow_dangerously_set_content: false
output_variable:
  description: The output will in json format.
````

User Intent Recognition: The user intent was recognized using Natural Language Understanding and Natural Language Processing to understand the user intent for the call. The model has performed relatively well with about ~96% accuracy to identify the different classes of intent in solving a classification problem. Below is the confusion matrix to showcase the results

Figure 7: User Intent Recognition Results

User Intent Recognition Results				
	Doctor Appointment	Appointment Cancellation	General Enquiry	SUM
Doctor Appointment	75 36.41%	0 0.00%	2 0.97%	77 97.40% 2.60%
Appointment Cancellation	1 0.49%	73 35.44%	2 0.97%	76 96.05% 3.95%
General Enquiry	1 0.49%	2 0.97%	50 24.27%	53 94.34% 5.66%
SUM	77 97.40% 2.60%	75 97.33% 2.67%	54 92.59% 7.41%	198 / 206 96.12% 3.88%

Figure 8: User Intent Recognition System Prompt

```yaml
name: RecognizeUserIntentPromptTemplate
template: |
  <message role="system">
   You an Intelligent Language Agent, which can recognize the user intentions from the user prompts.
   You are asked to analyze the user intent based on the user current message and user summarized chat history.

   To serve the user request, user intent should be recognized as one of the below
    - SchedulingDoctorAppointment - If the user is trying to schedule a doctor appointment.
    - CancellingDoctorAppointment - If the user is trying to cancel an existing doctor appointment.
    - GeneralEnquiry - If user is asking about certain general provider related information such as address, specialization etc.

   # Examples:
    - SchedulingDoctorAppointment = ["I am mandy, looking to book an appointment with doctor nancy"]
    - CancellingDoctorAppointment = ["Hi, I am john. Would like to cancel my upcoming reservation on 06/18/2025"]
    - GeneralEnquiry = ["Hi, I am trying to find doctors specialization of your practice", "What is the address of the doctors office?"]

   #Output Examples: The output you should produce should be an enum of one of the above intents. No comments, no other things.
    - Example Output 1:
       { "RecognizedIntent": "SchedulingDoctorAppointment"
  </message>
  <message role="user">
  {{userInput}}
  </message>
template_format: handlebars
description: This a user intent recognition prompt template.
input_variables:
  - name: userInput
    description: user message.
    is_required: true
    allow_dangerously_set_content: false
  - name: chatHistory
    description: user summarized chat history
output_variable:
  description: A json object with recognized user intent.
```

Question & Answer Extraction Results: Our novel transformative agent performed exceptionally well in extracting the answers to the questions that the LLM model was asked to extract via system prompt and RAG pattern. Below are some of the converted audio texts and extracted answers from the transcription by LLM model for different types of caller needs. When the system doesn't recognize the answers to the required questions based on the user intent. The system marks the questions as N/A and provides immediate feedback to the customer service agent on the call to get answers to the unanswered questions.

Table 1: Question and Answer extraction results

Converted audio text	**Question 1 & Extracted Answer**	**Question 2 & Answer**	**Question 3 & Answer**

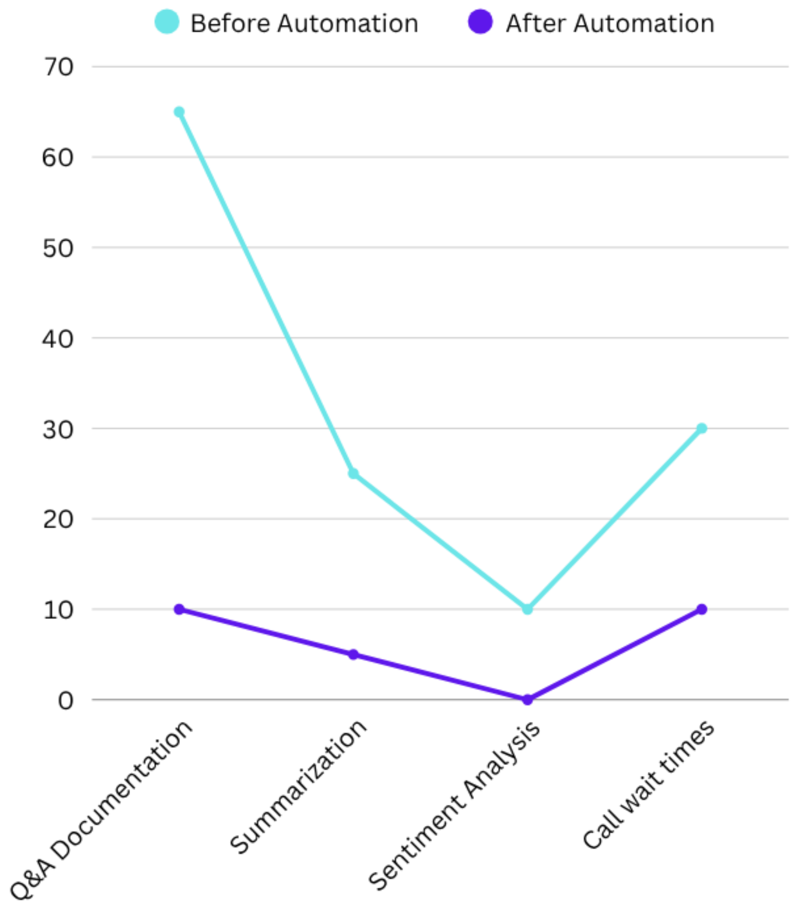
Hi, my name is John, I have had headache fever for the past three days. My age is 52.	What is the Patient name? Ans: John	What are the symptoms of patients? Ans: headache, fever	Number of days since the symptoms started and patient age. Ans: 3 days, 52 years
I am Rebecca, I am looking forward to scheduling an appointment with Doctor Jay for next Friday at 3PM	What is patient name? Ans: Rebecca	When is the patient looking for appointment? Ans: Next Friday 06/20/2025 at 3PM EST	Which doctor is the patient looking for? Ans: Jay
I want to cancel my appointment scheduled for next Tuesday.	What is patient name? Ans: N/A	When is the appointment scheduled? Ans: 06/17/2025	Which doctor is the appointment for? Reason behind the cancellation. Ans: N/A

Manual Task Automation Performance Results: Our transformative framework performed significantly well in automating various manual tasks of a healthcare customer service representative. Below are some of the results in percentages.

Table 2: Automation results

Task Type	**Time Spent Before Automation**	**Time spent after automation**
Question and answers documentation	65%	10%
Summarization	25%	5%
Sentiment Analysis	10%	0%
Call wait times	~15 – 30 mins	~5 – 10 mins

Figure 9: Automation Performance Results



7. Discussion:

This section of the article will discuss more the advantages, challenges and possible solutions of the currently proposed Artificial Intelligence agent framework powered by pre-trained LLM models for automating the Healthcare Compliance data during a live customer service call. In this section we will also be discussing the future of our agentic framework.

7.1 Challenges & Possible Solutions:

Using Pre-trained LLM models with "Prompt Engineering" is a very effective and efficient methodology for developing scalable and adaptable AI agents. Although the methodology is very adaptable, it comes with its own challenges such as

"Grounded Responses": LLM models are often known for their hallucinations as they are trained most general and predated data. They are trained to behave as per the consumer system requirements, making them generalized models. So, without the contextual and up to date system information, it's very hard for the model to provide grounded responses. One of the most widely used solutions to this problem is using Retrieval-Augmented Generation (RAG) pattern which allows us to inject the system specific contextual data into the system prompts for the LLM model to use and provide grounded responses.

Another solution is to control the imagination of the LLM models using "temperature" and "top" parameters.

"Token Limitation": Another well-known challenge of using LLM model is "token limitation". LLM models often suffer from token limitations where they cannot fetch the entire contextual data and inject into the system prompt for accurate grounded responses. This problem mainly arrives in the systems where the data volume is very heavy and the LLM model must access that data to perform the required operations. Some of the solutions to this problem are "Truncation" – truncating the beginning and ending of the sentences to reduce the number of tokens. "Chunking" – Breaking the data into smaller chunks and fed in batches to the LLM model to meet the token limitations. "Summarization" – sharing the summarized information rather than lengthy documents can reduce the token size. "Stop word Removal" – removal of stop words can also reduce the size of the token. "Vectorization" – Injecting vector data via vector databases can dramatically reduce the size of the tokens as the data is shared in embeddings.

"Accent Recognition": The main challenge of the speech to text conversion is to recognize the user accent to accurately predict and transcribe the audio content into the text file. The AI models often fail to produce viable results when the user accent cannot be recognized accurately. To overcome this challenge, this AI system implement deep Long Short term RNNs [10], [15] and advance speech recognition methodologies powered with deep neural networks and transformers to improve the accuracy of the speech to text process.

"Medical Terms Recognition" – Healthcare often has sophisticated terminology, which is hard to understand, unless provided with guidance. Our Agent framework overcomes this challenge by injecting the provider specific healthcare terminology via Plugins. Implementing RAG pattern to improve the accuracy of such terminology in the transcribed audio text.

7.2 Future work:

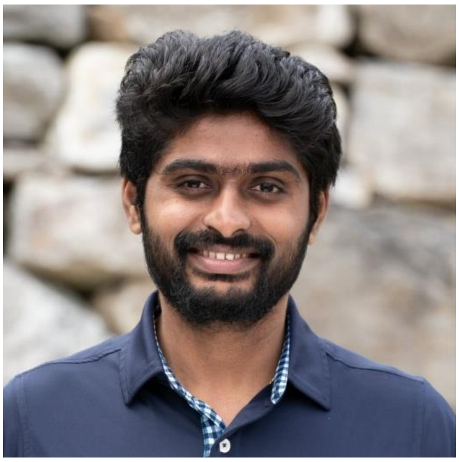
Our agent is specifically, created to work with English language not with any others. This limitation restricts wide adaptability of the system with different languages. Our target is to enable multi-language support for the agent for its speech to text, sentiment analysis and transcription features. Along with multi-lingual support our future work includes MCP server injection, multi-agent architecture for efficient scaling and Neural Voice generation for creating a empathic customer service agent system to reduce the cognitive load on the human workforce.

8. Conclusion.

Automating the tedious compliance and healthcare questionnaire provides the healthcare providers with greater flexibility, higher satisfaction rate and higher retention rates. Our novel transformative framework is expected to save about 75% of manual documentation time and save about 70% of expected revenue loss of the healthcare providers. The pre-trained LLM model provides our system with greater flexibility, higher adaptability and reduces the training costs and maintenance costs. Even with some of the well-known challenges of using pre-trained LLM models, the possible solutions and overall advantages outshine their shortcomings. Also, their rapid evolving nature makes the system more adaptable.

Finally, as discussed above in the Future work section of the article. Our framework is expected to address >~1.2 billion of revenue loss in the healthcare sector by boosting the performance of the customer care services. Our target is to expand the system to have multi-lingual support.

9. Authors' Biography



Leela Naga Sai Vamsi Krishna Dogiparthi, An Innovative Enterprise Applications Architect | Microsoft Certified Professional (MCP) | AI Engineer | trying to create novel solutions to hard problems using Artificial Intelligence.

References

1. "The Top Trends in Healthcare Contact Centers to Watch in 2025," Default. Accessed: Jun. 17, 2025. [Online]. Available: https://computer-talk.com/blogs/the-top-trends-in-healthcare-contact-centers-to-watch-in-2025
2. Uttiya, "6 Customer Support Trends and Predictions: What Your Business Needs to Know in 2025," Kommunicate Blog. Accessed: Jun. 14, 2025. [Online]. Available: https://www.kommunicate.io/blog/top-customer-support-trends-and-predictions/
3. R. T. Anderson, F. T. Camacho, and R. Balkrishnan, "Willing to wait?: The influence of patient wait time on satisfaction with primary care," BMC Health Serv. Res., vol. 7, no. 1, p. 31, Feb. 2007, doi: 10.1186/1472-6963-7-31.
4. C. Bleustein, D. B. Rothschild, A. Valen, E. Valatis, L. Schweitzer, and R. Jones, "Wait times, patient satisfaction scores, and the perception of care," Am. J. Manag. Care, vol. 20, no. 5, pp. 393–400, May 2014.
5. J. Sevilla, "Can AI Scaling Continue Through 2030?," Epoch AI. Accessed: May 01, 2025. [Online]. Available: https://epoch.ai/blog/can-ai-scaling-continue-through-2030
6. C. Giattino, E. Mathieu, V. Samborska, and M. Roser, "Artificial Intelligence," Our World Data, Sep. 2023, Accessed: May 01, 2025. [Online]. Available: https://ourworldindata.org/artificial-intelligence

7. M. Loukili, F. Messaoudi, and M. E. Ghazi, "Sentiment Analysis of Product Reviews for E-Commerce Recommendation based on Machine Learning".
8. S. S. Ranjani and S. Kalaichelvi, "Sentiment Analysis of Product Reviews for Ecommerce Recommendation based on Machine Learning," vol. 14, no. 5, 2025.
9. X. Fang and J. Zhan, "Sentiment analysis using product review data," J. Big Data, vol. 2, no. 1, p. 5, Dec. 2015, doi: 10.1186/s40537-015-0015-2.
10. A. Graves, A. Mohamed, and G. Hinton, "Speech Recognition with Deep Recurrent Neural Networks," Mar. 22, 2013, arXiv: arXiv:1303.5778. doi: 10.48550/arXiv.1303.5778.
11. "(PDF) Deep Learning Based Question Answering System: A Review," ResearchGate. Accessed: Jun. 14, 2025. [Online]. Available: https://www.researchgate.net/publication/348733305_Deep_Learning_Based_Question_Answering_System_A_Review
12. M. Negrão and P. Domingues, "SpeechToText: An open-source software for automatic detection and transcription of voice recordings in digital forensics," Forensic Sci. Int. Digit. Investig., vol. 38, p. 301223, Sep. 2021, doi: 10.1016/j.fsidi.2021.301223.
13. I. Aattouri, H. Mouncif, and M. Rida, "Call Center Customer Sentiment Analysis Using ML and NLP," in 2023 14th International Conference on Intelligent Systems: Theories and Applications (SITA), Nov. 2023, pp. 1–7. doi: 10.1109/SITA60746.2023.10373715.
14. "449_Sentiment.pdf." Accessed: May 27, 2025. [Online]. Available: https://www.ijirset.com/upload/2025/may/449_Sentiment.pdf
15. B. Subbulakshmi, M. N. Devi, M. Sivakumar, V. S. A, V. S. K, and K. M. T, "Efficient AI-Powered Audio-to-Text Transcription: A GUI-Enhanced Stack with EXE Build for Innovation in Communications," in 2023 International Conference on Energy, Materials and Communication Engineering (ICEMCE), Dec. 2023, pp. 1–7. doi: 10.1109/ICEMCE57940.2023.10434163.