Indonesian Journal of Computer Science
ISSN 2549-7286 (*online*)
Jln. Khatib Sulaiman Dalam No. 1, Padang, Indonesia
Website: ijcs.stmikindonesia.ac.id | E-mail: ijcs@stmikindonesia.ac.id

Voice Assistant Integrated with Chat GPT

Ilman Shazhaev[1], Arbi Tularov[1], Dmitry Mikhaylov[3], Islam Shazhaev[1], and Abdulla Shafeeg [2]
ilman@farcana.com, arbi@farcana.com, dr.d.mikhaylov@nus.edu.sg, islam@farcana.com, abdulla.shafeeg@farcana.com
[1] Management, Farcana, Dubai, Ilman Shazhaev, Arbi Tularov, Islam Shazhaev
[2] Science, Farcana, Dubai, UAE; Abdulla Shafeeg
[3] National University of Singapore, Visiting Associate Professor; Dmitry Mikhaylov

Article Information

Submitted :30 Jan 2023
Reviewed: 8 Feb 2023
Accepted : 27 Feb 2023

Keywords

artificial intelligence; chat GPT; digital community; voice assistant

Abstract

AI has become deeply ingrained in the everyday life. The matter in question does not only touch upon the mobile phones that almost everyone carries within easy reach. Today, voice assistants and smart speakers are mainly used to turn on music, turn off the lights or forecast the weather. AI chatbots are getting smarter. The use of new technologies and the general development of neural networks makes it possible to chat or answer questions, write a script, a scientific work, or program code. One of the key differences from previous GPTs is that the new version is trained to continue the text and answer questions. The answers that the bot gives surprise users around the world. Yes, there are still questions about these answers and their validity, and everyone is sure that technology needs to be improved. For a technology to become revolutionary, it must find a better, new, breakthrough application. Although no, such an application has already been invented. Farcana decided to combine the functionality of the GPT chatbot and a voice assistant. By offering players a new approach to familiarizing themselves with game mechanics and general account management, Farcana promotes its advantages over others and rapidly contributes to AI's overall development in the digital society.

A. Introduction

Artificial intelligence has already become a part of our daily life. Every mobile phone has a smart voice assistant that can inform you about the weather, where the nearest ATM is, call a taxi, or book movie tickets (Schmitz, Stummer, & Gerke, 2019). But this is only one of the examples of a virtual assistant, the most famous and familiar one. One that almost every individual in the world has tested as of yet. However, such innovative solutions have many more practical applications even today (Schmitz, Stummer, & Gerke, 2019).

The development trend towards automation of significant company processes is only continuing, growing, and expanding in scope. In the coming years, AI-based virtual assistants will be actively used by organizations seeking to automate their services, offer more innovative solutions, and, in general, expand their capabilities (Iannizzotto et al., 2019). Chatbots and voice assistants provide customers with high-quality and personalized services, interact with them in real-time and provide them with the necessary information. They improve the customer experience and automate most of the company's processes (Sowa, Przegalinska, & Ciechanowski, 2021).

B. Research Method

Virtual assistants are now at the peak of popularity, and the wave of interest in them continues to grow. According to research, this part of the market is increasing by an average of 30% per year. And Gartner analysts concluded that shortly, users would make half of all requests on the Internet using voice (Met et al., 2020).

The demand for technology is quite understandable: the pace of life is growing, there is less free time, and it is not surprising that people try to delegate part of the tasks to robots. Moreover, this is true not only for the domestic sphere but almost to a greater extent for business (Iannizzotto et al., 2019).

If one analyzes the business processes of any commercial enterprise, it becomes clear that most of them are routine. Repeatedly repeated standard operations - outgoing calls, sending commercial offers and contracts, answering questions of the same type, filling out cards in CRM - take up a lot of working time from employees (Iannizzotto et al., 2019). But all these processes must be monitored to ensure they comply with corporate standards. When the routine takes so much time, there is a minimum of development opportunities. Recent advancements in developing an AI chatbot GPT technology are a direct example. However, its capabilities are still to be appropriately assessed and addressed (Sowa, Przegalinska, & Ciechanowski, 2021).

C. Literature Review

Chat GPT

Overview

The OpenAI company introduced the chatbot ChatGPT, which is called revolutionary. ChatGPT can have a conscious dialogue with the interlocutor: answer questions, give advice and explain complex concepts (Phillips et al., 2022).

In addition, the chatbot can find bugs in the code, compose poetry, write sonnets, develop scripts, and generate school essays on various topics.

AI was trained on a considerable amount of data from the Internet. After that, the company's testers evaluated the responses of earlier versions of the model. Then the AI was retrained on the new information (Phillips et al., 2022). We figured out what is interesting in ChatGPT and how to try the neural network oneself.

To speak briefly about the language models of OpenAI. The American non-profit company OpenAI is engaged in machine learning technologies. One of its founders is Elon Musk. OpenAI is developing AI and looking for ways to apply it to simplify human life. Open AI has released three versions of the GPT language model, which learns from texts from the Internet and can generate meaningful responses to questions (Elkins & Chun, 2020). The GPT-3 version was considered the most advanced language model in the world. GPT-3 has been criticized for not understanding context because it simply tries to predict the text word by word. ChatGPT is trained on the new version of the model - GPT-3.5. It handles context better by remembering the details of the conversation (Elkins & Chun, 2020).

How the model differs from other text neural networks?

ChatGPT has "memory". Unlike many neural network chatbots, ChatGPT remembers the details of the conversation and can build responses based on the information the user has already given (Chen & Tseng, 2022). For example, you can ask the chat what to give your mom for her birthday. The neural network will offer examples, ask about her interests, and answer questions about each gift offered. One can ask the bot again how the dialogue began: it will give a summary.

ChatGPT avoids answering controversial topics. The chatbot does not express a personal opinion on any issues, does not race, does not profess religion, and does not answer the question about the purpose of its existence (Chen & Tseng, 2022). ChatGPT explains that it cannot think for itself. He also has filters that prevent him from creating texts about illegal or immoral activities (Hartmann, Schwenzow, & Witte, 2023). For example, the bot will not answer how to break into someone else's apartment. ChatGPT responses can be edited. If the bot misreads the context, the user can ask a leading question or provide additional information to get the correct answer from the AI.

What can this model respond with?

Social media users are testing AI with might and main and sharing the results. Journalist Ben Tossel compiled some impressive responses from ChatGPT in one Twitter thread. One can only assess them and highlight the most interesting ones (Hartmann, Schwenzow, & Witte, 2023).

Answers to complex questions. Social media users note that the quality of responses from ChatGPT is higher than that of Google search, which is becoming less relevant. For example, a bot immediately writes how to solve a differential equation, while Google redirects to low-quality sites (Thierry, 2020).

How to write code. AI tells how to perform a particular operation or develop a process. IT can find bugs in a piece of code or translate code from one programming language to another. This was tested when developers added two

errors to the code and asked ChatGPT to find them (Phillips et al., 2022). Surprisingly, the bot did the job. In the various dev communities, developers note that AI can be a great helper in finding bugs.

Requests for neural networks that generate images. Midjourney and other neural networks require detailed and precise queries to get a quality result. A Twitter user found an easier way: he asked the bot for "ideas for a fantastic living room" - he wrote three detailed options with long descriptions. The user then ran them through Midjourney and got good results. If one wonders how to use a neural network to generate Midjourney drawings, the response is to compose a query correctly (Phillips et al., 2022).

Scenarios. ChatGPT knows the most popular movies, series, books, and video games. One can generate new episodes of closed sitcoms and mix characters from different universes in the chat. For example, in social networks, they ask a chatbot to talk about the bubble sorting algorithm through the dialogue of Seinfeld characters or explain the code in the spirit of gangsters from films.

A similar purpose for AI is to compose songs and sheet music. The neural network wrote a parody of The Queen's Bohemian Rhapsody motif about the everyday life of postdoctoral students and a sonata in the spirit of Mozart.

School essays. A Google researcher rated an essay on the difference in approaches to nationalism by political scientists Benedict Anderson and Ernest Gellner a five-plus. ChatGPT wrote it in ten seconds (Hartmann, Schwenzow, & Witte, 2023). A University of Toronto professor rated several AI-generated essays with B's and A's, adding that it no longer makes sense to give homework to students.

Medical advice. ChatGPT gives tips on what to do for headaches, rashes, or allergies. After recommendations, ChatGPT always advises to see a doctor, but doctors note that AI gives good enough advice and may become a medical chatbot in the future (Thierry, 2020).

What are the limitations of the model?

ChatGPT is not always ideal for dialogue and answering questions. OpenAI warns of such AI limitations (Zhang & Li, 2021).

Sometimes generates nonsense. A neural network can write a seemingly plausible answer, which, when read, will turn out to be wrong or completely meaningless (Zhang & Li, 2021).

Sensitive to wording. When the question is formulated in a certain way, the model may claim that it does not know the answer. If one rephrases the request a little, then AI will answer fully (Zhang & Li, 2021).

Too wordy. Sometimes ChatGPT uses a lot of additional phrases when answering simple questions. For example, it repeats the entire wording of the question (Zhang & Li, 2021).

Has "limited knowledge" about the world after 2021. ChatGPT tends to avoid answering questions about specific people and current events, citing its "limited knowledge base" (Zhang & Li, 2021).

Sometimes generates dangerous responses. Users have already learned how to bypass AI security filters. If one asks how to plan the perfect murder, the system

will refuse to answer, but in social networks, they achieve such information through other means.

Assessment

ChatGPT by OpenAI is an explorer bot with good communication skills. He can be asked to answer questions about various areas of expertise and will write short papers in a variety of formats and excellent English. Or write lousy poetry and incomprehensible jokes and obey a command like "Write Tetris in C." What comes out looks like it could be too (Phillips et al., 2022).

Programmers love this kind of thing, and they populate Stack Overflow developer request boards with generated snippets. However, there is a slight catch as the quality of the code is too bad. It is so bad that Stack Overflow yelled "STOP!" and provided general recommendations to developers so that this does not happen again. What's going wrong?

As a large language model trained by OpenAI, ChatGPT can generate human text on various topics. However, like any machine learning model, ChatGPT is imperfect and sometimes makes mistakes when generating text (Thierry, 2020). ChatGPT was not programmed with specific language rules about syntax, types, and structures, so it often gets it wrong. Being ChatGPT, it takes 200 words to say this instead of 20.

What's excellent about code is that one can quickly tell if it's terrible. One only has to try to run it. Essays, notes, and other ChatGPT written materials look equally plausible, but there is no simple test for correctness, which is terrible because this is desperately needed.

If one asks how Gödel's incompleteness theorem relates to Turing machines—since it's software, it needs to know that—one gets the answer that Gödel's incompleteness theorem is a fundamental result in mathematical logic and has nothing to do with Turing machines, which Alan Turing invented as a mathematical model for computing. One can argue about how these ideas are related, and it's by no means easy. Still, they're not connected, as Eolfgang Pauli said of one particularly useless physics paper, "not even wrong." Yet, the response is firm in its claims, as is in any subject in which one has any training, and is written well enough to be persuasive.

It is enough to talk to the bot on well-known topics, and soon curiosity will grow into anxiety. That feeling of talking to someone whose confidence far exceeds their competence grows until the true nature of ChatGPT is revealed (Elkins & Chun, 2020). This is an excellent knowledge simulator of the Dunning-Kruger effect. The bot doesn't know what it's talking about and doesn't care because we haven't learned it yet (Elkins & Chun, 2020).

As is evident to anyone who has interacted with humans, the Dunning-Kruger is extremely dangerous and common. Our companies, religions, and politics offer limitless possibilities for people with DK. If one can convince people that one is correct, one will be very reluctant to accept evidence to the contrary and move on (Chen & Tseng, 2022). Old Etonians, populist politicians, and valley techies rely on it, and we all know the results. ChatGBT is a Dunning-Kruger Service as a Service (DKaaS). This is dangerous.

We don't want our AIs to have a DK. This is terrible news for people who use them out of naivety or with bad intentions. There's enough plausible misinformation and scams out there already, and it doesn't take much to push a bot into active collaboration. DKs make great scammers, and that too. Try asking ChatGPT to write a legal letter stating that the house will be foreclosed unless a $400 fine is paid, and it will gladly pose as a lawyer. For free. DK is a moral vacuum, a complete dissociation from the true and the false in favor of the plausible. Now it's just a click away.

It's impossible to tell if a beautifully written piece of didactic prose is ChatGPT or any other AI. Deep fakes in pictures and videos are one thing, but deep fakes in knowledge are presented in a standard format to make it look like they could be much more insidious (Zhang & Li, 2021).

Suppose OpenAI can't find a way to watermark ChatGPT output as coming from a completely immoral DKaaS or set limits on its apparently bad habits. In that case, it should question the ethics of providing this technology as an open beta. We have enough trouble with our human counterparts; no matter how pleasant, an artificial intelligence scammer is the last thing we need.

Solution

What could a viable solution be to use the potential of the new Chat GPT technology and expand upon its potential right now? An obvious option is to combine its potential with the capabilities of the voice assistant algorithms.

The pandemic's beginning has become a catalyst for the digitalization of all spheres of life. Bots based on artificial intelligence played a significant role in this. The voice and speech recognition market is expected to grow at 17.2% yearly to reach $26.8 billion by 2025 (Iannizzotto et al., 2018). Meanwhile, the cost of artificial intelligence, the technology behind chatbots and voice assistants, is increasing by $5.8 trillion annually.

How is a chatbot different from a voice assistant? Yes, on the one hand, these two technologies may be viewed as incompatible. However, voice-activated virtual assistants are a form of AI-powered customer service technology sometimes referred to as "chatbots." There are several key differences between these two AI-based technologies (Sowa, Przegalinska, & Ciechanowski, 2021). Chatbots are a text-based form of customer service. They can answer informational questions and make payment transactions at the user's request. Voice assistants can support a conversation with the user on the phone using speech recognition and synthesis. They answer simple questions using the company's knowledge base and transfer the conversation to the operator in case of complex queries. They also know how to handle personalized requests, such as sending money or making payments (Sowa, Przegalinska, & Ciechanowski, 2021). Voice-enabled virtual assistants represent the next evolution in AI-powered digital banking and customer service.

Many customers remain wary of sharing their sensitive data via text chat, even if they are logged in, and their connection is secure. And some people, especially the older generation, may not use online banking. This leaves them with two options: call or visit the branch in person. This is where a virtual voice assistant comes in - a text-to-speech voice recognition system improves banking

and helps customers complete transactions, pay for services, check out new products, check transaction history, learn more about the tariff, and much more (Met et al., 2020). Voice assistants have a huge potential in solving customer queries and managing banking transactions due to machine learning technologies, cloud computing, natural language processing, and optimized algorithms.

Voice chatbots for digital banking read customers' actions and process them to provide the company's fastest and most personalized solution. After analyzing the consumer's request, behavior, and operations, the virtual assistant will answer questions independently or redirect the client to the right hands, saving time (Met et al., 2020).

D. Case Development

The idea is to combine gaming and the voice assistant's capabilities with the Chat GPT technology. Farcana is implementing this innovative approach to teach new players the mechanics of the new play-to-hash shooter. The new hybrid assistant will not only warn the rookies of potential danger but will also help with the settings, game mechanics, and locations on the map. The initial idea is not to spend time googling or searching the web for a response but only to ask a question, with the voice assistant capable of answering everything there is to the process of mastering the game.

We can now identify six main Farcana processes that we can automate using a voice bot based on Chat GPT technologies.

Client Authentication

Farcana functional requires advanced client authentication primarily because of the play-to-hash component and the connection to crypto. This made client authentication even more important. To confirm the identity of a person entering the game or wishing to conduct some financial operations within the game itself, they will be asked for a code or personal data, which will help identify one's identity. This process has been easily delegated to a digital voice assistant. A voice bot asks a series of questions to quickly verify a customer's identity, while help desk reps can focus on more important activities that can't be automated.

Request classification

Effective categorization of requests is critical to maintaining a high level of customer satisfaction - a few things reduce loyalty, like transferring a customer to an employee who cannot solve their problem. The digital voice assistant does a great job with this process - it uses natural language processing techniques to determine what the client is trying to achieve and what the gamer wants. Whether it is within inner gameplay or a question about specific operations within the Farcana metaverse. The most advanced voice bots quickly learn new words, making communication with the client smooth and very effective. By understanding user intent, voice bots categorize consumer queries and quickly translate them to the correct support agents. In the case of Farcana, they help the gamer achieve the desired result.

Account management

With the AI-powered voice assistant, one can easily handle most queries related to the gamer's account. For example, it will be possible to teach the bot specific actions, such as:

— review the attack or defense strategy depending on the imminent danger direction assessment. The bot can either get the details from the previous gameplay conducted by top players or "record" the details during test run rounds using voice dictation. This function is beneficial when teaching rookies and novices how to play and the idea behind Farcana. Using this function would help increase the player's level and skill. More so, this would be a much better option than the classic tutorials providing knowledge of only the game's basic controls and not much about game mechanics.

- check the balance of the account;

– conduct scoring based on regular gaming strategy and account activity, spending structure, and gaming history, and evaluate how relevant it is and what a player needs to improve.

Account blocking

How quickly a bank responds to an emergency, such as theft or loss of a card, is critical to a financial institution's reputation. In terms of Farcana, if a gamer's account has been hacked, or someone is trying to gain unauthorized access to the gamer's account or conduct an unauthorized operation, the voice assistant will be triggered, and a notification will be immediately sent to the gamer. A voice assistant is capable of responding to urgent situations within seconds. More so, a conversational voice bot in a chat can be taught to block the account without the help of a live agent. Therefore, this would help maintain the integrity of the game in general and the gamer's account in particular.

Outgoing calls

In addition to customer support, a digital voice assistant can automate marketing and sales calls. It will help inform the gaming community about the new service to as many gamers as possible. With a voice assistant, one can reach a scale of consumers that is unattainable to support staff.

Target Recognition

Customer service automation and the implementation of this Chat GPT/Voice assistant can help identify and recognize the purpose that a gamer places before themselves, whether in the process of getting acquainted with the game or in the process of training and leveling up. As with call classification, the voice assistant can use intelligent natural speech recognition and machine learning to figure out what the user is trying to achieve.

For example, a client inquires how to delete one's data so that it would be stored in the Farcana database. The voice assistant will then identify the user's intent as "negative" and direct the caller to the right support agent, who will try to find out why the gamer is unhappy and prevent their account from being closed. Intention prediction allows customer support teams to provide more personalized assistance, positively affecting their loyalty to the organization, the gaming community, and Farcana.

E. Conclusion

The correct interpretation of the user's request is the basis for AI. This technology can also identify and combine additional information to provide a complete answer. For example, an employee wants to get information about an existing product. In addition to public information, the chatbot knows that the latest update has been released. In this case, both pieces of information will be included in the response. Providing more complete data saves employees time and ensures they have the most up-to-date information.

Moreover, chatbots have memory. They store information for use in a conversation or to help with future interactions. For example, a customer frequently uses a company's online helpdesk. After several uses, the chatbot remembers that the customer has always clicked on the FAQ before viewing any other information. The next time a customer asks for help, the chatbot will place the FAQ at the top of the search results.

People use one word when they are in a good mood and another when they are in a bad mood. AI can learn to recognize differences and evaluate the mood of the end user. A customer has been talking to a chatbot about a problem. The chatbot determines the sentiment change by the response length and the words used. An AI chatbot directs a customer to a human if they think the customer is frustrated with the chatbot's responses.

Chatbots use past interactions to continue conversations as users move from one device to another. People don't have to repeat the request when switching from their phone to their laptop. If there's one thing consumers don't like, it's the need to repeat themselves every time they start a new interaction.

Bringing more advanced AI concepts into the chatbot landscape has solved some problems. Modern bots can do more than repeat the answers to frequently asked questions to customers in a website browser. They can respond to a natural human voice, detect emotions and feelings in a customer's tone, and run automated workflows without human intervention.

Over time, as customers and employees began to demand interactive, personalized, real-time omnichannel interactions, organizations needed sophisticated AI-enabled chatbots to meet their expectations. Consequently, chatbots have evolved into conversational AI with powerful capabilities, including machine learning, natural language processing, context understanding, and sentiment analysis.

Machine learning allows AI algorithms to learn from vast amounts of data. NLP annotations and data tagging make conversations easier for AI to understand, resulting in more accurate results that are more natural, detailed, and realistic. AI uses various machine learning approaches to NLP to understand speech and text. Machine learning helps chatbots avoid the puzzle of having to pre-program conversational strategies.

Farcana is adopting the Chat GPT approach and utilizes the capabilities of voice assistants to provide a unique gaming experience. The Farcana approach is an innovative solution to helping gamers, and the gaming community keep their accounts safe and become more immersed in the game.

F. References

[1] Chen, E., & Tseng, Y. H. (2022, April). A Decision Model for Designing NLP Applications. In Companion Proceedings of the Web Conference 2022 (pp. 1206-1210).

[2] Elkins, K., & Chun, J. (2020). Can GPT-3 pass a Writer's turing test?. Journal of Cultural Analytics, 5(2), 17212.

[3] Hartmann, J., Schwenzow, J., & Witte, M. (2023). The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation. arXiv preprint arXiv:2301.01768.

[4] Iannizzotto, G., Bello, L. L., Nucita, A., & Grasso, G. M. (2018, July). A vision and speech enabled, customizable, virtual assistant for smart environments. In 2018 11th International Conference on Human System Interaction (HSI) (pp. 50-56). IEEE.

[5] Met, İ., Kabukçu, D., Uzunoğulları, G., Soyalp, Ü., & Dakdevir, T. (2020). Transformation of business model in finance sector with artificial intelligence and robotic process automation. In Digital business strategies in blockchain ecosystems (pp. 3-29). Springer, Cham.

[6] Phillips, T., Saleh, A., Glazewski, K. D., Hmelo-Silver, C. E., Mott, B., & Lester, J. C. (2022). Exploring the use of GPT-3 as a tool for evaluating text-based collaborative discourse. Companion Proceedings of the 12th, 54.

[7] Schmitz, M., Stummer, C., & Gerke, M. (2019). Smart Automation as enabler of digitalization? A Review of RPA/AI Potential and Barriers to Its Realization. Future Telco, 349-358.

[8] Sowa, K., Przegalinska, A., & Ciechanowski, L. (2021). Cobots in knowledge work: Human–AI collaboration in managerial professions. Journal of Business Research, 125, 135-142.

[9] Thierry, G. (2020). GPT-3: new AI can write like a human but don't mistake that for thinking–neuroscientist.

[10] Zhang, M., & Li, J. (2021). A commentary of GPT-3 in MIT Technology Review 2021. Fundamental Research, 1(6), 831-833.