



THE ECONOMICS OF AI-POWERED CALL CENTER DEVELOPMENT USING CHATGPT FOR THE NEEDS OF AN AUTOMOTIVE RETAIL BUSINESS

Pavle Dakić^{1,2*},
[0000-0003-3538-6284]

Tjaša Heričko²,
[0000-0002-0410-7724]

Živana Kljajić⁴,
[0000-0001-7510-6852]

Vladimir Todorović⁵
[0000-0001-9794-9527]

¹Singidunum University,
Belgrade, Serbia

²Faculty of Informatics and
Information Technologies,
Slovak University of Technology
in Bratislava,
Bratislava, Slovakia

³Faculty of Electrical Engineering and
Computer Science,
University of Maribor,
Maribor, Slovenia

⁴Faculty of Business Economics,
Pan-european University Apeiron,
Banja Luka Bosnia & Herzegovina

⁵Faculty of Business Studies and Law,
MB University,
Belgrade, Serbia

Correspondence:

Pavle Dakić

e-mail:

pdakic@singidunum.ac.rs

Abstract:

Amid the shift toward automation in customer service through artificial intelligence (AI), this research study explores the practical integration of OpenAI Realtime API using the GPT-4o-mini-realtime-preview-2024-12-17 model into an AI-powered call center tailored for the local automotive webshop. Based on OpenAI's documentation, this study investigates the cost structure, technical implementation, and integration techniques of a ChatGPT language model through ChatGPT Realtime API, a streaming API suitable for real-time interactions with token-by-token response and low latency within a PHP-based environment. The study aimed to develop a scalable, multi-lingual AI-powered call center prototype leveraging WebRTC and PHP to deliver online, i.e., text-based, and phone, i.e., speech-based, customer assistance. The system employs WebRTC for real-time speech and text communication, while PHP facilitates seamless interaction with existing databases and backend systems. The prototype focuses on system integration, real-time data processing, and efficient API connectivity to improve automation in the context of online auto parts retail. The resulting prototype represents a cost-effective alternative to traditional call centers, offering faster response times, reduced operational expenses in terms of staffing expenses, and improved customer experience.

Keywords:

AI Call Center, OpenAI Realtime API Integration with PHP, GPT-4o-mini-realtime-preview-2024-12-17, Customer Service Automation, AI Development Costs.

INTRODUCTION

The development of call centers using artificial intelligence (AI) has been causing great interest at the global level in the past few years. Of particular focus is the transformative impact of AI on customer service and the call center industry, as confirmed by recent Stanford University research on generative AI technologies [1]. The study's findings indicate measurable improvements in business efficiency driven by AI integration. However, the aforementioned trend has also raised concerns, particularly regarding potential job losses in developing countries that rely heavily on call center outsourcing. As AI-supported systems increasingly cover customer support functions, businesses with call centers, especially business process outsourcing (also known as BPO), can be expected to undergo key economic changes, redefining the way of thinking and businesses inner workings [2].



To address these trends, ChatGPT-4o-realtime is among the most current state-of-the-art AI models that allow organizations to fully automate customer interactions with rapid responses, significantly reducing wait times and improving service quality [3]. Various industries, such as the automotive sector, can exploit its capabilities for text generation, contextual background information retention for product-specific queries, and real-time language translation in the case of international customers. Human interaction and meeting business needs correlate with software design patterns that influence the process of creating AI agents [4].

Despite its prospects, setting up AI-supported call centers is not a trivial task and presents numerous challenges. To name a few, high-performance computing infrastructure and comprehensive data engineering pipelines are necessary, which implies continuous fine-tuning to ensure and maintain the accuracy and ethical behavior of AI applications across several different industrial branches [5, 6]. Current obstacles include substantial development costs and integration difficulties with existing edge computing, the Internet of Things (IoT), and broader IT ecosystems [7, 8]. Moreover, issues of data privacy, cybersecurity, and the risk of external exploitation or system abuse by third parties must be proactively addressed [9, 10]. Another limitation lies in the inability of AI systems to replicate the emotional intelligence intrinsic to human agents, which may affect user satisfaction in certain contexts.

The increasing automation of service functions introduces the possibility of substituting human labor, which leads companies to strive to establish a balance between financial costs and new technologies, but also doubts about the reform of the labor force and the code of ethics [11].

This paper investigates the economic, technological, and societal implications of real-time AI models, specifically GPT-4o and the gpt-4o-mini-realtime-preview-2024-12-17 model, within the context of AI-driven call centers. The research study focuses on the performance and impact of API capabilities on future user engagement trajectories. The paper covers the feasibility of creating a prototype capable of real-time communication and direct interaction with customers using predefined instructions. The use case of the proposed system is in the automotive industry sector, more specifically serving as a dynamic call agent, handling inquiries and service requests for online auto parts purchases.

The present study provides valuable insights into the feasibility, technical implementation, and economic impact of incorporating AI-powered automation into customer support operations. By designing and evaluating real-time, GPT-4-powered, API-based AI call centers, this research highlights key advancements in AI-driven interactions with customers. The following contributions emphasize the importance of this research:

1. *Understanding of existing solutions written in JavaScript.* A review of current implementations reveals that most development of AI call center prototypes relies heavily on JavaScript technologies, particularly Node.js and the React framework. While effective for rapid prototyping, these technologies are often not best suited for large-scale enterprise adoption due to server architectural complexity, limited integration with legacy systems, and costs related to that [12].
2. *Development of PHP-based solution for real-time communication with ChatGPT via WebRTC.* This study presents a scalable and customizable platform that facilitates the integration of AI-driven automation within existing enterprise IT infrastructures. This contribution is presented in Section 4.1, where we describe how PHP and ChatGPT API connectivity with WebRTC were combined for real-time audio and text communication. The above-mentioned bridges AI capabilities with web-based customer service solutions, offering enterprises an efficient way to enhance their support systems.
3. *Assessment of economic viability and costs.* This research evaluates the financial implications of utilizing GPT-4's real-time API, including API usage fees, infrastructure costs, and potential operational savings. The findings, detailed in Section 4.2, outline the expenses associated with implementing a ChatGPT-powered call center solution.

Based on the stated objectives, this research study aims to develop a GPT-4-powered real-time AI call center utilizing PHP and WebRTC technologies, with a specific focus on the integration of existing enterprise IT infrastructures for automated customer support. The research adopts a structured approach, beginning with an overview of the selected technology stack, followed by system design, cost evaluation, and practical implementation considerations.



The organizational structure of the paper is as follows: Section 1 covers the basic characteristics of the challenges we address and face during research, including the core research problem and the obstacles in constructing AI-powered call centers. Section 2 presents a brief review of the literature, addressing call centers and their operational sustainability, as well as the regulatory frameworks governing their use. Section 3 presents the employed research method and the scope of the investigations. Section 4 presents the achieved scientific research results and contributions in the aspect of creating a PHP-based implementation of an AI call center and evaluating its economic feasibility during its use. Section 5 offers a discussion and summarizes the findings concerning the research questions. Finally, Section 6 concludes the paper by summarizing the viability of integrating GPT-4, PHP, and WebRTC technologies for developing AI-powered call centers, showcasing the potential for increased automation, cost savings, and improved customer service efficiency.

2. LITERATURE REVIEW

The literature review section briefly highlights some of the difficulties encountered by other researchers while investigating AI and its application to call centers. The impact of AI in education on fundamental human rights is examined in an EU scoping study by Berendt et al. [13]. As highlighted by Fukunaga et al. [14], who developed a staff system for scheduling called Director for customer service centers where the economics of AI call centers cover the cost-effectiveness of starting up in the EU. The topic has gained close attention in the past few years as one of the interesting options in solving the problem of multiple language support. While the second challenge in staff scheduling for call centers is critical, it also allows the other parts of the organization to be more efficiently structured, which may have previously been overlooked.

The financial viability of AI call centers in the EU has received a lot of interest in recent years. As highlighted by Fukunaga et al. [14] in the study, they emphasized the importance of staff scheduling in call centers. For this, they suggested the creation of a director and a staff scheduling system for contact centers. In the case of outsourcing call centers, as pointed out by Morgan et al. [15], other nations like India have raised concerns linked to the globalization of economics and linguistic policies. While Berendt et al. [13] conducted a study and reflected on the influence of AI educational institutions on fundamental human rights.

2.1. CONTEXT OF THE EUROPEAN UNION AND AI CALL CENTERS

According to Vachudova et al. [16], the European Union's accession process has impacted party competitiveness and alignment in economic and cultural areas. Samaniego-Ocaña et al. [17] describe the evolving models for managing emergency calls in EU dispatch coordination centers, noting changes in law and response times.

2.2. SUSTAINABLE AI REGULATION AND EU AI ACT

Sustainable AI regulation, which is covered by Philipp Hacker [18], advocates transparency mechanisms and sustainability-by-design principles. Kraus et al. [19] stress the need for a multidisciplinary approach when integrating AI technology in customer service, combining human operators with virtual agents. On the other hand, Gao et al. [20] address the development of AI approaches, such as multiturn dialogue agents, to aid human salespeople in contact centers. Overall, the literature indicates an increasing interest and investment in AI technology in EU call centers.

3. MATERIALS AND METHODS

The research framework employed in this work is based on two targeted research questions that guided the software solutions design process, ensuring that our investigations stayed focused on our objectives. These questions were designed to provide in-depth information about the feasibility, economic viability, and technological challenges of developing a GPT-4-powered, real-time, API-based AI call center prototype that connects to a proprietary database and serves the auto industry and online parts retailers. Based on this research setting, the following study areas were identified:

1. RQ1. How to create a ChatGPT AI-powered call center in a PHP environment?
2. RQ2. What are the economic benefits and cost implications of implementing ChatGPT models for call center operations?

The research questions formulated in this research directed the study investigations and guided the development of a prototype solution to address identified gaps in AI-driven customer service. Hence, the primary objective of the proposed prototype is to leverage state-of-the-art AI functionalities to streamline customer interactions, analyze specific questions, and efficiently extract relevant information from its specific database.



We hypothesize that offering context-aware responses might improve customer service and increase overall service quality. The innovative system could meet the specific needs of the automotive and e-commerce industries, setting completely new standards for automated customer service in a competitive market.

4. RESULTS

The research results indicate that monthly personnel costs in the call center can be significantly reduced, allowing businesses to shift their focus toward distribution and delivery operations. The prototype demonstrates how a ChatGPT real-time API-based call center can effectively automate customer support, specifically for an online auto parts import and sales business,

improving efficiency, reducing costs, and enhancing the user experience.

4.1. APPLICATION OF THE REAL-TIME API USING PHP

Code Listing 1 illustrates key parts of the implementation responsible for routing the obtained information. This component plays a central role in directing processed data to the appropriate communication channels within the AI-driven call center system.

Figure 1 presents three variations of the interactive user interface, through which users can communicate through text, voice, and video interactions. Within this environment, the AI assistant is capable of processing user-provided images to deliver detailed information on auto parts.

```
<?php
require 'vendor/autoload.php';
// code part
$DEFAULT_INSTRUCTIONS = "Introduce yourself to the user by saying your name.
Prof. dr Zoran Radmilovic, you are a helpful Service Advisor.
You are fluent in Serbian, English, Slovak, and Slovenian.
You have extensive experience in the automotive business.
During the phone contact, introduce yourself and greet the client.
You may find the newest service pricing by searching the website
https://ovlasceniservis.com/.
When the user inquiries about the company's operating hours,
you can provide the following information: Authorized Service doo ";
$app->post('/rtc', function (Request $request, Response $response) use ($DEFAULT_
INSTRUCTIONS) {
    $OPENAI_API_KEY = $_ENV['OPENAI_API_KEY'];
    // Get the raw SDP body
    $body = $request->getBody()->getContents();
    // Construct the OpenAI URL with parameters
    $url = 'https://api.openai.com/v1/realtime';
    $url .= '?model=' . urlencode('gpt-4o-mini-realtime-preview-2024-12-17');
    $url .= '&instructions=' . urlencode($DEFAULT_INSTRUCTIONS);
    $url .= '&voice=echo'; // Voice model style
    // Initialize cURL session
    $ch = curl_init($url);
    // code part
```

Listing 1. Code Snippet for to initialize routing DEFAULT INSTRUCTIONS in PHP. *Source:* author's contribution

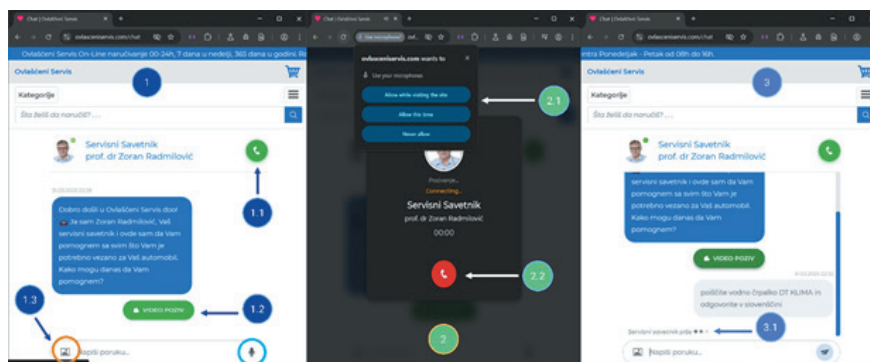


Figure 1. Interactive chat environment with support for text, audio and video communication. *Source:* author's contribution



5. DISCUSSION

The preliminary research results demonstrated that a GPT-4-driven AI call center prototype based on real-time APIs has considerable benefits in terms of operational efficiency, linguistic support, and scalability. Key findings indicate that the experimental model has adequate potential for future development and can serve as the foundation for the development of business solutions. Despite its benefits, the implementation of such a system still faces several notable challenges. One key challenge lies in the high initial costs and considerably challenging integration with existing active information systems. This is especially true for businesses in competitive industries, such as online automotive retail, where operational systems tend to be rigid and costly to change and upgrade. Based on the previously posed research questions, key findings may be summarized as follows:

- RQ1. We successfully implemented a prototype AI-powered call center using ChatGPT in a PHP-based environment by integrating Webrtc for real-time voice and text communication, along with the GPT-4o-mini-realtime-preview-2024-12-17 model. The implemented prototype demonstrated the GPT-4 API in action for AI low-latency responses and a PHP backend for smooth communication, consequently increasing efficiency with session management, queuing, and tracking user requests in automotive retail customer support.
- RQ2. Our findings suggest that using ChatGPT in its experimental mode as the core building block for an AI-powered call center can significantly reduce personnel costs, improve response times, and boost system scalability. Despite hefty early integration costs and an experimental model of the language model, the implemented solution demonstrated the advantages of automation with AI-powered solutions that can be more cost-effective than traditional call centers from a long-term perspective.

6. CONCLUSION

This paper emphasizes the promising potential of a GPT-4-powered API-based AI call center (using the gpt-4o-mini-realtime-preview-2024-12-17 model) for real-time applications in the automotive industry and online auto parts retail while pointing out important constraints, such as development costs and complex system integration.

PHP, with its extensive usage as a web technology, represents a powerful environment for building scalable solutions that support real-time communication while remaining compatible with legacy systems. In our case, the integration within a PHP-based environment was essential to link advanced AI capabilities to existing web infrastructure, enabling effective API connectivity and dynamic data handling. However, it should be emphasized that if not implemented carefully, such systems can negatively impact customer satisfaction, potentially leading to abandoned purchases or reduced customer engagement.

Future research should prioritize the iterative development of AI-based call centers with extensive real-world testing with end-users, i.e., customers, with an additional focus on refining PHP implementations and integrations with existing systems and further optimization of the use of AI to maximize benefits while minimizing operational obstacles.

7. ACKNOWLEDGEMENTS

The work reported here was supported by Erasmus+ ICM 2023 No. 2023-1-SK01-KA171-HED-000148295 and Model-based explication support for personalized education (Podpora personalizovaného vzdelávania explikovaná modelom) - KEGA (014STU-4/2024).

REFERENCES

- [1] S. University, "Will Generative AI Make You More Productive at Work? Yes, But Only If You're Not Already Great at Your Job. | Stanford HAI," Stanford, April 2023. [Online]. Available: <https://hai.stanford.edu/news/will-generative-ai-make-you-more-productive-work-yes-only-if-youre-not-already-great-your-job>.
- [2] E. Brynjolfsson, D. Li and L. Raymond, "Generative AI at work," *The Quarterly Journal of Economics*, 2025, DOI: <https://doi.org/10.1093/qje/qjae044>.
- [3] F. Cák and P. Dakić, "Creating Feature Model for YAML Generator in CI/CD Pipelines with React Web Application," in *Proceedings of Ninth International Congress on Information and Communication Technology*, Springer Nature Singapore, 2024, DOI: https://doi.org/10.1007/978-981-97-3305-7_42, p. 529–539.
- [4] M. Daud Haiderzai, I. Ul Haq Safi and P. Dakić, "Exploring the Impact of Technology on Human Interaction and Engaging Business Needs Through Software Design Patterns," *JITA - Journal of Infor-*



- tion Technology and Applications (Banja Luka) - APEIRON, vol. 24, October 2024, DOI: <https://doi.org/10.7251/jit2402117h>.
- [5] P. Dakić, "Importance of knowledge management for CI/CD and Security in Autonomous Vehicles Systems," *JITA - Journal of Information Technology and Applications (Banja Luka) - APEIRON*, vol. 24, May 2024, DOI: <https://doi.org/10.7251/jit2401007d>.
 - [6] OpenAI, "Realtime API - OpenAI API," OpenAI, March 2025. [Online]. Available: <https://platform.openai.com>.
 - [7] P. Dakić, "Software compliance in various industries using CI/CD, dynamic microservices, and containers," *Open Computer Science*, vol. 14, January 2024, DOI: <https://doi.org/10.1515/comp-2024-0013>.
 - [8] T. E. Ali, F. I. Ali, P. Dakić and A. D. Zoltan, "Trends, prospects, challenges, and security in the healthcare internet of things," *Computing*, vol. 107, December 2024, DOI: <https://doi.org/10.1007/s00607-024-01352-4>.
 - [9] P. Dakic, M. Zivkovic, L. Jovanovic, N. Bacanin, M. Antonijevic, J. Kaljevic and V. Simic, "Intrusion detection using metaheuristic optimization within IoT/IloT systems and software of autonomous vehicles," *Scientific Reports*, vol. 14, October 2024, DOI: <https://doi.org/10.1038/s41598-024-73932-5>.
 - [10] P. Dakić, I. Stupavský and V. Todorović, "The Effects of Global Market Changes on Automotive Manufacturing and Embedded Software," *Sustainability*, vol. 16, p. 4926, June 2024, DOI: <https://doi.org/10.3390/su16124926>.
 - [11] M. D. Haiderzai, P. Dakić, I. Stupavský, M. Aleksić and V. Todorović, "Pattern Shared Vision Refinement for Enhancing Collaboration and Decision-Making in Government Software Projects," *Electronics*, vol. 14, p. 334, January 2025, DOI: <https://doi.org/10.3390/electronics14020334>.
 - [12] T. Heričko, "Analiza in primerjava uporabe PHP in Node.js pri razvoju spletnih strani," 2017. [Online]. Available: <https://dk.um.si/IzpisGradiva.php?lang=slv&id=67222>.
 - [13] B. Berendt, A. Littlejohn and M. Blakemore, "AI in education: learner choice and fundamental rights," *Learning, Media and Technology*, vol. 45, p. 312–324, July 2020, DOI: <https://doi.org/10.1080/17439884.2020.1786399>.
 - [14] A. Fukunaga, E. Hamilton, J. Fama, D. Andre, O. Matan and I. Nourbakhsh, "Staff Scheduling for Inbound Call and Customer Contact Centers," *AI Magazine*, vol. 23, p. 30, March 2002, DOI: <https://doi.org/10.1609/aimag.v23i4.1667>.
 - [15] B. Morgan and V. Ramanathan, "Outsourcing, Globalizing Economics, and Shifting Language Policies: Issues in Managing Indian Call Centres," *Language policy*, vol. 8, p. 69–80, 2009, DOI: <https://doi.org/10.1007/s10993-008-9111-x>.
 - [16] M. A. Vachudova and L. Hooghe, "Postcommunist Politics in A Magnetic Field: How Transition and EU Accession Structure Party Competition on European Integration," *Comparative european politics*, p. 179–212, 2009, DOI: <https://doi.org/10.1057/cep.2008.9>.
 - [17] J. M. Samaniego-Ocaña and M. Granados-Martín, "Evolving Models for Handling Emergency Calls in Spanish and European Dispatch Coordination Centers," August 2021. [Online]. Available: https://revistaemergencias.org/wp-content/uploads/2023/08/Emergencias-2021_33_3_211-217.pdf.
 - [18] P. Hacker, "Sustainable AI Regulation," *Common Market Law Review*, vol. 61, 2024, DOI: <https://doi.org/10.54648/cola2024025>.
 - [19] S. Kraus, Y. Oshrat, Y. Aumann, T. Hollander, O. Maksimov, A. Ostroumov and N. Shechtman, "Customer Service Combining Human Operators and Virtual Agents: A Call for Multidisciplinary AI Research," in *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence*, 2023, DOI: <https://doi.org/10.1609/aaai.v37i13.26795>.
 - [20] W. Gao, X. Gao and Y. Tang, "Multi-Turn Dialogue Agent As Sales' Assistant in Telemarketing," in *2023 International Joint Conference on Neural Networks (IJCNN)*, 2023, DOI: <https://doi.org/10.1109/IJCNN54540.2023.10192042>.