World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

Check for updates

AI-enhanced conversational avatars for immersive interaction

Etibar Aliyev [1,*] Oluyinka Adedokun [2], Adetutu Temitope Fabusoro [3] and Babatunde Keshinro [4]

[1] *Informatics and Analytics, UNC Greensboro, Greensboro, NC, USA.*
[2] *Industrial & Systems Engineering and Engineering Management, University of Alabama in Huntsville, IL, USA.*
[3] *Education Policy Organization and Leadership, University of Illinois, Urbana Champaign, IL, USA.*
[4] *Industrial and Systems Engineering, North Carolina A&T State University, Greensboro, NC. USA.*

World Journal of Advanced Research and Reviews, 2025, 26(01), 627-634

Publication history: Received on 25 February 2025; revised on 03 April 2025; accepted on 05 April 2025

Article DOI: https://doi.org/10.30574/wjarr.2025.26.1.1101

Abstract

This paper reviews the significance of key technologies of AI-enhanced conversational avatars and their applications for immersive interaction, including increased AI system adoption, enhanced trustworthiness, and personalized and engaging interactions. Using empirical data and case studies, this paper provides evidence regarding recent advancements that facilitate high-fidelity and style-controllable avatar generation, providing customizable options for users for online social interactions. This is possible by integrating several AI modules like text-to-speech, speech recognition, and open-domain dialogue systems for responsive and realistic virtual characters. Considering technologies like natural language processing (NLP), affective computing, and multimodal interaction are examined, the framework's effectiveness for AI avatars in conversational agents and its constituent elements like appearance, nonverbal and verbal communication, machine learning capabilities, and appearance characteristics are demonstrated. Although numerous benefits are derivable from the integration, experts have cautioned about discrimination, biases, and potential psychological effects related to AI avatars, hence it is critical to ensure responsible development and interaction.

Keywords: AI-enhanced conversational avatars; Artificial intelligence; Conversational avatars; Immersive interaction

1. Introduction

AI-enhanced conversational avatars are evolving as powerful tools for immersive interaction through combined advanced AI technologies and human-based representations. According to [1], the evolution of conversational agents (CAs) is marked by strides of technological advancements from agelong traditional models to modern-day generative systems. These agents have been applied in multiple domains like healthcare, education, and marketing, and their development has been characterized by distinct research waves which show theoretical paradigms and technical evolutions [2]. Through recent advancements in large language models and statistical computing, more domain-agnostic deployments and natural interactions have become more possible with even far-reaching benefits [3]. Integrating CAs with immersive technologies like virtual reality (VR) offers great opportunities, especially in promoting user trust via expressiveness, responsiveness, and embodiment [4].

As these developments are shaping future applications and research regarding conversational agents, AI-enhanced conversational avatars as digital representations of AI systems closely resembling humans foster human-like conversations and interactions [5]. The avatars comprise key elements like appearance characteristics, nonverbal cues, verbal communication, and machine learning capabilities [6]. The advantages of using these avatars include increased AI system adoption, enhanced trustworthiness, and personalized engagement. Recent advancements in audio-driven technology and AIGC encourage customizable avatars using different conversational abilities and visual styles [7].

* Corresponding author: Etibar Aliyev

Moreover, they demonstrate abilities to transform user experience across industries. AI-based avatars are used in educational settings for contextualized and individualized instruction through synthetic voices and large language models for enhanced user interactions [8].

This paper discusses the role of AI-enhanced conversational avatars for immersive interaction. To ensure a comprehensive understanding of the subject, the objectives of this review paper are:

- To examine key artificial intelligence (AI) technologies that enable conversational avatars
- To demonstrate the effectiveness of AI avatars for higher engagement and personalization
- To analyze the integration of AI modules in developing realistic and responsive virtual characters for immersive interactions
- To emphasize the importance of interaction frameworks and responsible development for AI-enhanced conversational avatars while evaluating the potential risks associated

2. Application of AI-Enhanced Conversational Avatars for Immersive Interaction

2.1. Technologies Enabling AI-Enhanced Conversational Avatars

Advancements in machine learning, speech recognition, and natural language processing (NLP) are common consequences of AI-revolutionized human-machine communication [9]. AI-driven conversational tools like chatbots and virtual assistants have enhanced efficiency and user engagement across multiple domains. According to [6], the core elements of AI avatars for conversational agents range from appearance characteristics to verbal and nonverbal aspects, and machine learning. These systems use pattern-based, deep learning and machine learning approaches to simulate human behavior while incorporating sentiments and emotions [10]. While challenges remain in real-time processing, ethical concerns, and contextual understanding, ongoing research has focused on multimodal and empathetic AI systems development to bridge the human expectations and machine capabilities gap [9].

In addition, the increasing use of NLP and deep learning proves the ability for photoreal face synthesis, where advanced AI-synthesized avatars facilitate virtual AI companions and real-time deepfakes [11] [12]. For affective and more natural virtual humans, multi-layered sensing and AI systems have been developed for voice stress analysis, recognizing facial emotion, and sensing the environment [13]. Essentially, these systems encourage empathetic conversations with sensed humans through gestures, facial expressions, and voice-based interactions. In the past, the focus has been on assembling expressive facial animation systems from components in the public which demonstrates text-to-speech synthesis applications with simple conversational agents and expression control.

Furthermore, recent speech technology advancement enables more expressive and natural interactions between AI-enhanced conversational avatars and humans. Recognition and speech synthesis systems are being developed to complement multimodal interactions in augmented and virtual reality environments, fostering more immersive experiences [14]. There are also continuous efforts to identify ways to incorporate expressivity and emotion into artificial voice technologies through post-processing techniques and markup tags to modulate synthesized speech [15]. With various implications of these developments in different fields including artificial intelligence, human-computer interaction, and augmentative communication, expressive speech recognition systems can distinguish between different emotions through affective robot-child communication, while new speech synthesizers have the potential to produce expressive and high-quality nonsense speech using recorded databases and prosody transplantations [16].

2.2. Applications of AI-Enhanced Conversational Avatars

According to Bibliometric analyses, GPT-3 has diverse applications that transcend computer science, especially as chatbots are emerging as common use cases in healthcare, education, and customer service. The evolution of AI-enhanced conversational avatars has propelled telemedicine consultations, automated tutoring, and intelligent virtual assistants in global businesses. Similarly, chatbots continually help to streamline processes and improve user engagement [17].

2.2.1. Case Study 1: Virtual Tutors and Immersive Simulations

Integrating AI-driven avatars in education has transformed students' learning experiences through personalized instructions, real-time feedback, and the development of immersive simulations for hands-on training [18]. Specifically, Duolingo's AI-powered virtual tutor uses machine learning algorithms to adapt to users' language proficiency while personalizing lesson plans. Through this, proficiency scores are improved in learners using AI-powered tutors

compared to their peers using traditional learning platforms. It also led to increased user engagement with higher retention rates due to the personalized feedback and interactive lessons features.

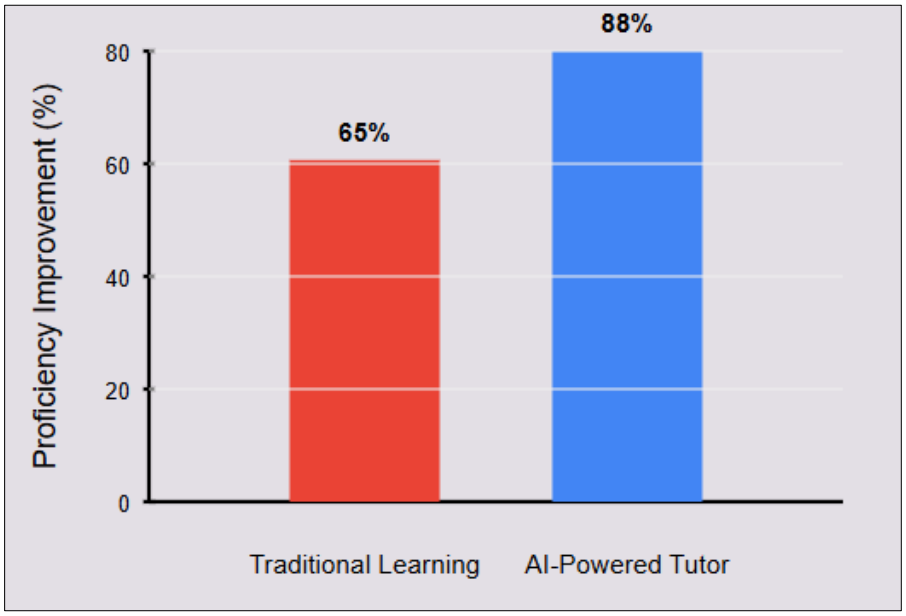

Figure 1 Impact of Duolingo's AI-Powered Virtual Tutor on Learning Outcomes

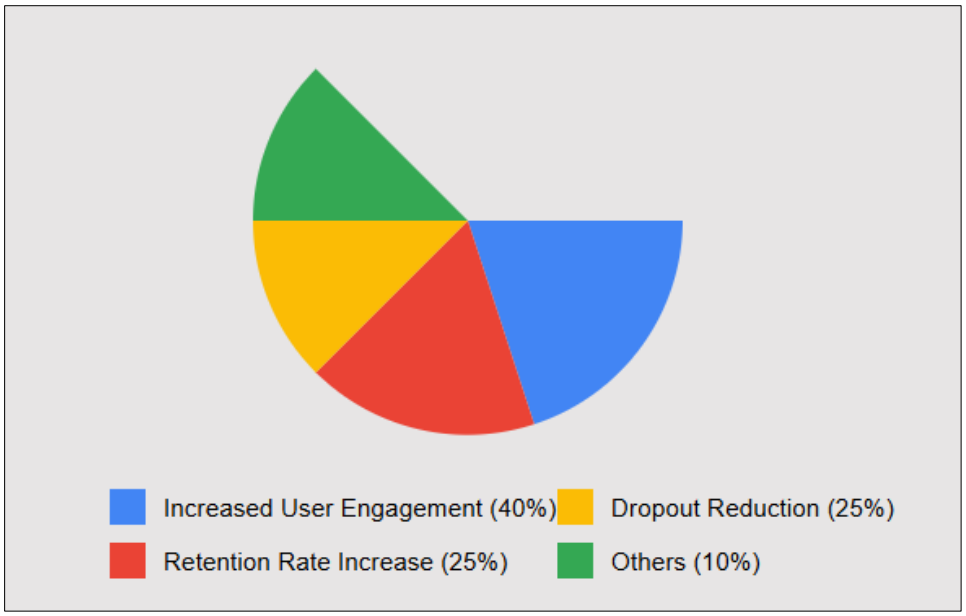

Figure 2 Impact of AI Tutors on Learners' Engagement and Retention

2.2.2. Case Study 2: Impact of OpenAI's GPT on Efficient System

While recent studies have explored the use and application of OpenAI's GPT models across multiple domains, the case in Zimbabwe's tech industry has striking interests, where ChatGPT has shown potential to improve programming support and web development, influencing information, system, and service quality [19]. Integrating AI-driven conversational avatars led to a resultant improvement in the accuracy of responses, responsiveness & reliability, and user satisfaction & engagement levels.

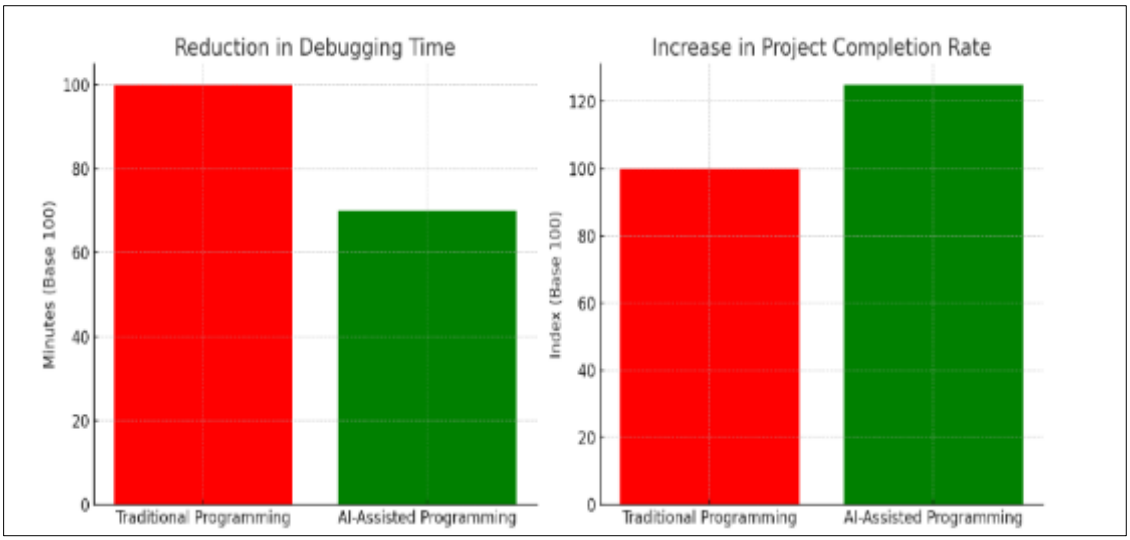

Figure 3 Impact of ChatGPT across IT sectors in Zimbabwe

2.2.3. Case Study 3

ChatGPT's performance was evaluated with conversational QA corpora and BERT similarity scores, providing insights into its strengths, limitations, and potential areas for improvement [20]. The study showed that ChatGPT has a high accuracy (85%) in structured queries. In unstructured and complex conversations, however, it had lower performance (65%). This indicates room for enhanced contextual awareness.

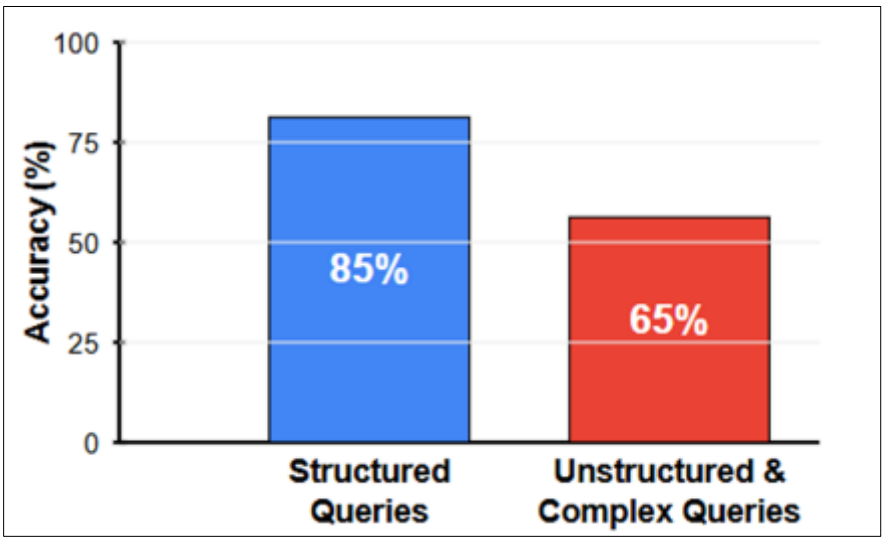

Figure 4 ChatGPT Performance Evaluation Based on Conversational QA Corpora

2.2.4. Case Study 4: Virtual Therapists and Patient Support

Woebot is an AI-powered mental health chatbot which has revolutionized access to psychological support. As a virtual therapist, the bot provides 24/7 mental health assistance, emotional support, and cognitive behavioural therapy (CBT) techniques using conversational interfaces. In a 2021 clinical study, experts found that Woebot users witnessed a 30% reduction in depression symptoms only after two weeks of use, while more than 80% of users reported improved feelings of emotional support after their interaction with AI-powered therapy bots [21]. Mental health chatbots provide up to 50-60% daily engagement rates while adhering to self-help programs. 60% of users also reported the use of mental health chatbots for emotional support at least 3 times a week.

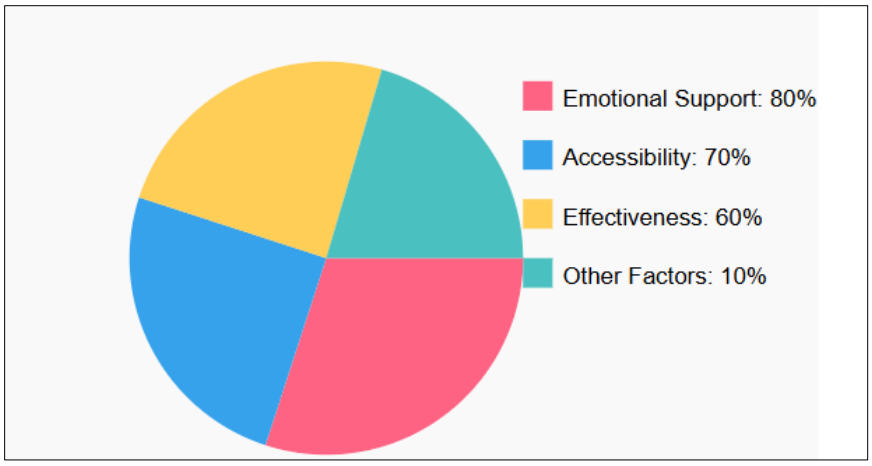

Figure 5 User satisfaction with AI mental health chatbots [Woebot AI-Powered Mental Health Chatbot]

3. Challenges and Limitations

AI-enhanced conversational avatars face some ethical and technical challenges in immersive interactions. Some common hurdles are real-time responsiveness and scalability, especially for edge devices showing computational constraints [22]. Latency issues and computational requirements constitute significant obstacles to effective, seamless interaction, while ethical concerns often arise due to the creation of realistic virtual environments and responsive avatars, reshaping social interactions and human experiences [23]. During the collection of precise motion tracking data, sensitive psychological and physical information can be revealed, which raises concerns about potential misuse and data security [24].

According to [25], challenges like authenticity and cultural adaptability in language learning applications remain. Due to the characteristic feature of multiple dimensions of the design space of conversational agent avatars impacting user interaction and perception, careful consideration is required for contextual and task-specific aspects [26]. Integrated XR and AI technologies lead to the creation of responsive avatars and realistic virtual environments which may reshape human experiences and interactions, including reflection of biases and preferences from personal space in virtual reality (VR), which makes interaction data a significant privacy issue. Biases in algorithm design, data collection, and decision-making processes can result in unfair outcomes for the individual user and society. These biases often derive from prejudices related to those shown by humans as seen in Microsoft's Taybot [27].

Moreover, although emotionally intelligent chatbots can enhance user experiences via artificial empathy, there are challenges regarding data security and empathetic responses [28]. Further research would be required to address these limitations and stimulate their effectiveness in various settings [25]. They also highlight the need for ethical considerations while developing AI avatars to address data protection, emotional manipulation issues, and others related to the long-term social consequences of embodied artificial intelligence in virtual worlds [24]. Similarly, the increased use of frameworks to measure and reduce bias in AI chatbots, especially the Chatbot Bias Assessment Framework and other counter-stereotypic imagining approaches can play fundamental roles [27].

4. Future Directions and Research Opportunities

Due to the recent advancements in affective computing and multimodal emotion recognition, new avenues have been opened for effective and increased human-computer interaction [29]. Brain-computer interfaces (BCIs), neurophysiological modalities, and behavioral signals have potential for emotion recognition [30]. Similarly, although significant progress has been achieved in enabling AI systems to identify and express effects, more research is required to explore the impact of these affective predictions on machine understanding of cognitive states and human social behaviors [3]. As ubiquitous emotion recognition using data fusion from several multimodal, mobile devices emerges, integrating numerous modalities like eye contact, gestures, and facial expressions is critical for optimized human-computer interaction [31-37]. Therefore, future research should focus on overcoming multimodal interaction challenges while exploring its technological advancement potential in brain-computer and emotional AI interfaces.

Not only that, personalization and adaptation for sustainable user engagement models present significant research opportunities. Personalized lifelong learning alongside augmented cognition can leverage lifelong user modelling,

especially with learners controlling their user models by themselves [38-42]. Through this, the human-centred and technical challenges of privacy and control can be sufficiently addressed. According to [33], personalization is essential in human-robot interaction to maintain engagement and build trust for longer periods, especially in applications regarding elderly care, educational robots, and companion robots. Focusing research on this area to sustain interaction quality and relationship-building for long-term, effective interactions is recommended [34]. As user modelling, personalization and adaption methods evolve to resolve potential issues, future research should also explore the development of privacy-preserving, robust lifelong user models to improve long-term engagements in human-robot interactions and refine relevant techniques for use in pervasive computing environments [39-43].

5. Conclusion

This review discusses notable progress in AI-powered conversational avatars supported by strong evidence from several case studies. Duolingo's AI tutors achieved better learning outcomes and increased retention through tailored teaching methods, whereas ChatGPT reached 85% accuracy with structured questions but faced challenges in handling intricate dialogues. Woebot users saw their depression symptoms decrease by 30% after two weeks while 80% reported better emotional support and 60% used the platform three times per week.

AI avatars revolutionize human-computer interactions by implementing technologies that support context-aware and emotionally intelligent communication. The combination of natural language processing affective computing and multimodal systems has produced virtual characters that detect emotions and respond suitably transforming industries from education to mental healthcare.

In sum, the necessity to tackle various challenges persists with real-time responsiveness data security algorithmic bias and ethical concerns about emotional manipulation as key issues. The exploration of future research avenues encompasses the development of multimodal emotion recognition systems alongside brain-computer interfaces and privacy-preserving personalization models to maintain sustained user engagement. The ongoing evolution of these technologies will make AI-enhanced conversational avatars merge digital and physical interactions into indistinguishable experiences which demand attention to technological capabilities alongside ethical considerations.

References

[1] S. Schobel, A. Schmitt, D. Benner, M. Saqr, A. Janson and J. M. Leimeister, "Charting the Evolution and Future of Conversational Agents: A Research Agenda Along Five Waves and New Frontiers," Information Systems Frontiers, vol. 26, pp. 729-754, 2024.

[2] A. Alnefaie, S. Singh, B. Kocaballi and M. Prasad, "An Overview of Conversational Agent: Applications, Challenges and Future Directions," Proceedings of the 17th International Conference on Web Information Systems and Technologies.

[3] V. Mathur and A. Singh, "The Rapidly Changing Landscape of Conversational Agents," ArXiv, 2018.

[4] C. Yildrim, "An Immersive Model of User Trust in Conversational Agents in Virtual Reality," in Third International Conference on Transdisciplinary AI (TransAI), 2021.

[5] Y. Liu and K. L. Siau, "Human-AI Interaction and AI Avatars," in Interaccion, 2023.

[6] P. Khampuong, P. Nilsook and P. Wannapiroon, "Artificial Intelligence Avatar for Conversational Agent," in 2023 Research, Invention, and Innovation Congress: Innovative Electricals and Electronics (RI2c), 2023.

[7] Y. Guo, R. Yan, Y. Wu and S. Ma, "Styleself: Style-Controllable High-Fidelity Conversational Virtual Avatars Generation," in 2024 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), 2024.

[8] M. C. Fink, S. A. Robinson and B. Ertl, "AI-Based Avatars Are Changing the Way We Learn and Teach: Benefits and Challenges," Frontiers in Education, 2024.

[9] O. M. A. Naser, H. A. Faraj and K. M. Elmbark, "AI-Driven Conversation Interfaces: Advancements in Human-Machine Communication," International Journal for Research in Applied Science & Engineering Technology (IJRSAT), vol. 12, no. 12, 2024.

[10] S. Kusal, Patil, S. A, J. Choudrie, K. V. Kotecha, S. Mishra and A. Abraham, "AI-Based Conversational Agents: A Scoping Review from Technologies to Future Directions," IEEE Access, 2022.

[11] M. Allouch, A. Azaria and R. Azoulay, "Conversational Agents: Goals, Technologies, Vision and Challenges," Sensors, 2021.

[12] Z. Wang, K. Nagano, H. Li, L. Hu, L. Goldwhite, H. Kung, A. Agarwal, L. Wei, Y. Chen, Q. Xu, J. Seo and H. Luo, "AI-Synthesized Avatars: From Real-Time Deepfakes to Virtual AI Companions," ACM SIGGRAPH 2020 Real-Time Live, 2020.

[13] S. DiPaola and O. Yalcin, "A Multi-Layer Artificial Intelligence and Sensing Based Affective Conversational Embodied Agent," in 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), 2019.

[14] Y. Wang, L. Wang and K. L. Siau, "Human-Centered Interaction in Virtual Worlds: A New Era of Generative Artificial Intelligence and Metaverse," International Journal of Human-Computer Interaction, vol. 41, no. 2, pp. 1459-1501, 2025.

[15] L. Orynbay, R. Bibigul, P. P, M. Blaz and E. Ziga, "Recent Advances in Synthesis and Interaction of Speech, Text, and Vision," Electronics, vol. 13, no. 9, p. 1726, 2024.

[16] E. Szekely, "Expressive Speech Synthesis in Human Interaction," University College Dublin, 2015.

[17] C. Varghese, E. M. Harrison, G. O'Grady and E. J. Topol, "Artificial Intelligence in Surgery," Nature Medicine, vol. 30, no. 5, pp. 1257-1268, 2024.

[18] B. N. Shirazi, A. A. Safavi, E. Aftabi and G. Salimi, "The Integration of Virtual Reality and Artificial Intelligence in Educational Paradigms," in 2024 11th International and the 17th National Conference on E-Learning and E-Teaching (ICeLeT), 2024.

[19] A. Sibanda, "Customer Adoption of ChatGPT for Web Development and Programming Assistance in the Zimbabwe Tech Industry," in International Student Conference on Business, Education, Economics, Accounting, and Management (ISC-BEAM), 2024.

[20] A. Rangapur and H. Wang, "ChatGPT-Crawler: Find Out if ChatGPT Really Knows What It's Talking About," ArXiv, 2023.

[21] P. L. Yeh, W. C. Kuo, B. L. Tseng and S. Y. H, "Does the AI-Driven Chatbot Work? Effectiveness of the Woebot App in Reducing Anxiety and Depression in Group Counseling Courses and Student Acceptance of Technological Aids," Current Psychology, pp. 1-13, 2025.

[22] A. Basit and M. Shafique, "tinyDigiClones: A Multi-Modal LLM-Based Framework for Edge-optimized Personalized Avatars," in 2024 International Joint Conference on Neural Networks (IJCNN), 2024.

[23] E. Craig, "Ethical Challenges in the Convergence of XR and AI," 2020.

[24] L. E. Buck and R. McDonnell, "Security and Privacy in the Metaverse: The Threat of the Digital Human.," 2022.

[25] N. H. Nguyen and V. P. Pham, "AI Chatbots for Language Practices," International Journal of AI in Language Education, 2024.

[26] M. Rashik, M. Jasim, K. Kucher, A. Sarvghad and N. Mahyar, "Beyond Text and Speech in Conversational Agents: Mapping the Design Space of Avatars," in Proceedings of the 2024 ACM Designing Interactive Systems Conference, 2024.

[27] H. Beattie, L. A. Watkins, W. H. Robinson, A. Rubin and S. Watkins, "Measuring and Mitigating Bias in AI-Chatbots.," in 2022 IEEE International Conference on Assured Autonomy (ICAA), 2022.

[28] M. Rostami and S. Navabinejad, "Artificial Empathy: User Experiences With Emotionally Intelligent Chatbots," AI and Tech in Behavioral and Social Sciences, vol. 1, no. 3, pp. 19-27, 2023.

[29] O. O. Fagbo, O. B. Adewusi, D. A. Atakora, T. S. Lawrence, S. A. Olufemi and Z. Ezevillo, "Designing Intelligent Cyber Threat Detection Systems Through Quantum Computing," 2025.

[30] Z. He, Z. Li, F. Yang, L. Wang, J. Li, C. Zhou and J. Pan, "Advances in Multimodal Emotion Recognition Based on Brain-Computer Interfaces," Brain Sciences, vol. 10, 2020.

[31] I. J. Jacob, "Harnessing Power of Multimodal Interaction, their Challenges and Future Prospect – A Review," December 2023.

[32] J. Zheng, S. Qiu, C. Shi and Q. Ma, "Towards Lifelong Learning of Large Language Models: A Survey," ACM Computing Surveys, vol. 57, no. 8, pp. 1-35, 2025.

[33] B. Irfan, A. Ramachandran, S. Spaulding, D. F. Glas, I. Leite and K. L. Koay, "Personalization in Long-Term Human-Robot Interaction," in 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2019.

[34] T. S. Lawrence, P. Oyirinnaya and A. A. Adesola, "The Crucial Role of Artificial Intelligence in Fintech for Suptech and Regtech Supervision in Banking and Financial Organizations," International Journal of Artificial Intelligence Research and Development (IJAIRD), vol. 3, no. 1, pp. 38-50, 2025.

[35] Azeez, M., Nenebi, C. T., Hammed, V., Asiam, L. K., & James, E. (2024). Developing intelligent cyber threat detection systems through quantum computing. International Journal of Science and Research Archive, 12(2), 1297-1307.

[36] V. G. Dimitrova, T. Kuflik, D. N. Chin, F. Ricci, P. Dolog and G. Houben, "User Modeling, Adaptation, and Personalization," 2014.

[37] Azeez, M., Ugiagbe, U. O., Albert-Sogules, I., Olawore, S., Hammed, V., Odeyemi, E., & Obielu, F. S. (2024). Quantum AI for cybersecurity in financial supply chains: Enhancing cryptography using random security generators. World Journal of Advanced Research and Reviews, 23(1), 2443-2451.

[38] Alli, Y. A., Bamisaye, A., Bamidele, M. O., Etafo, N. O., CHKIRIDA, S., Lawal, A., ... & Al Nageim, H. (2024). Transforming waste to wealth: Harnessing carbon dioxide for sustainable solutions. Results in Surfaces and Interfaces, 100321.

[39] Fagbo, O. O., Adewusi, O. B., Atakora, D. A., Lawrence, T. S., Olufemi, S. A., & Ezevillo, Z. (2025). Designing intelligent cyber threat detection systems through quantum computing.

[40] Hammed, V., Bankole, A. A., Akinrotimi, O., & Ayanleye, O. (2024). Silver nanoparticles (AGNPs): A review on properties and behavior of silver at the nanoscale level. International Journal of Science and Research Archive, 12(2), 1267-1272.

[41] CHanson, E., Nwakile, C., & Hammed, V. O. (2024). Carbon Capture, Utilization, and Storage (CCUS) Technologies: Evaluating the Effectiveness of Advanced CCUS Solutions for Reducing CO2 Emissions. Results in Surfaces and Interfaces, 100381

[42] Hammed, V., Eyo, D. E., Omoloja, T. O., Kolawole, M. I., Adeyemi, A., & Tolulope, A. K. (2024). A review of quantum materials for advancement in nanotechnology and materials science.

[43] Keshinro B. (2022). Image Detection and Classification: A Machine Learning Approach. Available at SSRN 4281011.