Petzy: AI-powered interactive pet companion

Sanjana M[1], *Shreya* P[1*], Dr. *Ravikumar* G K[1], and Dr. *Chaitra* HV[1]
[1]BGS College of Engineering and Technology, Mahalakshmipuram, Bengaluru, Karnataka, India

Abstract. Existing literature on virtual agents shows strong progress in conversational AI but still lacks emotionally responsive, multilingual, and context-adaptive systems capable of sustaining long-term engagement across varied domains. Most current models remain text-centric and do not support real-time emotional reciprocation or cross-commodity use. To address this gap, the study presents Petzy, a 3D AI-powered virtual pet designed for friendly, low-pressure interaction. Its dual-zone behaviour design—Fun Zone and Intellectual Zone—supports emotional computing and adaptive engagement. Petzy identifies user emotions through speech and language analysis and responds using a multilingual intelligent chatbot paired with dynamic behavioral traits such as affection, vitality, thirst, and happiness. The system is built using React JS, Three JS, Blender for 3D modelling, Firebase for secure cloud storage, and FastAPI for backend routing, with Groq API and Ollama enabling high-performance LLM inference and multilingual speech detection. Experimental trials assessed latency, emotional accuracy, adaptive learning, and engagement. Results show Petzy delivers low response times, stable emotional feedback, and high user participation while functioning as both an emotionally aware companion and a domain-specific simulator for training and learning contexts.

1 Introduction

Artificial intelligence is having more and more of an effect on how people act, beyond just automating and computing. It is changing how people feel and interact with each other. People today want more than just chatbots that work; they want things that can understand them, change to fit their needs, and be friends. Traditional virtual assistant architectures, limited by scripted responses, frequently lack emotional awareness, leading to disengagement over time. Additionally, virtual training settings within health care, education, as well as professional skills development often depend on static either predetermined characters without contextual emotional depth, which reduces realism and pedagogical value. Petzy fills these gaps by adding sentimental analysis and adaptive dialogue to an interactive 3D space. It uses sentimental conversational artificial intelligence as well as voice synthesis to create a realistic experience with an animated pet model. The identical architecture can also create personas that are emotionally responsive and specific to a profession for training based on simulation. Petzy is different from text-only companions because it combines vision, speech, as well as gamified interaction to make an optimal

* Corresponding author: 04shreyap@gmail.com

platform for learning, entertainment, as well as medical or training use. The framework shows how smart systems can build emotional connections with users while still being easy to use, customizable, and keeping their data private.

2 Problem Statement

Most conversational agents don't have a lot of emotional continuity, don't remember context well, and don't include enough languages. Most virtual pet apps today rely on pre-made animations or scripted responses that don't change based on how the user feels or their cultural background. Simulated interactions employed in health care, education, as well as vocational training similarly lack contextual emotional depth; avatars are frequently pre-scripted rather than genuinely adaptive, diminishing the fidelity and educational value of simulation exercises. Also, emotional modeling in regular chatbots is mostly limited to English and a small number of emotional states, which makes them less inclusive and less interesting to users. To get around these problems, Petzy was made as an adaptable framework that combines emotion recognition from both voice and text inputs. This lets animations change based on the tone of the conversation. The system allows people to talk to each other in eighteen various dialects, including Sanskrit and several Indic scripts. This makes it easier for people from different cultures to use the system. Petzy's architecture goes beyond just being a friend. It lets you make simulation environments that can change based on the situation. For example, it can make simulated patients for students in medicine or interactive discussions participants for language learners through integrating multi-language Natural Language Processing in emotional state modeling. In addition, Petzy uses gaming-based engagement measurements (nutrition, playing, grooming, teleporting) that are connected to emotional well-being indicators. It also uses Firebase Firestore to make sure that users can access their data on any device without any problems. Petzy turns regular virtual animals and scripted simulators into efficient, personalized learning companions by combining emotional intelligence, simulation ability, and game-based motivational logic.

3 Related Work

1.1 Gamified Learning Chatbots. Benner et al. [1]

This research work shows that gamification makes learners more motivated, but it doesn't often include empathy. Petzy improves this by connecting psychological reactions scores with XP progression.

1.2 Gamification and Engagement. Hidayatulloh et al. [2]

This research work demonstrated enhanced engagement via rewarding systems; Petzy now expands that with persistent conduct tracking and personalized streaks.

1.3 Value Creation in Gamified Chatbot Interactions and Its Impact on Brand Engagement [3]

This research work highlighted a necessity for context-specific sentimental adaptation ; Petzy now reinforces this within public conversations through sentiment-aware LLM responses.

1.4 Emotionally Intelligent Systems. Heidrich [4]

This research work emotionally intelligent systems analyze tone, facial expressions, and linguistic cues to respond empathetically. Although impactful in education and healthcare, many lack real-time adaptation, underscoring the need for emotionally aware AI companions like Petzy

1.5 Virtual Pets and Avatars. Ramsay et al. [5]

This research work demonstrated that the protective conduct encourages a connection; Petzy translates such as emotional observation into concrete happiness indices.

1.6 Personalization in AI Companions. Pearson [6]

This research work stated dynamic mental shaping Petzy operates a continuous preferences log that customizes tone, illustration hue, and activity suggestions.

1.7 Speech Recognition and Emotion Detection. Basireddy [7] and Rastogi et al. [8]

This research work emphasized prosodic signals as sentiment indicators; Petzy integrates the Web speech API detection with tone-based response modulation.

1.8 Design and Implementation of Chatbots. Kumar et al. [9]

This research work discussed recent chatbot architectures, including rule-based vs AI-based bots, and suggested the integration of DialogFlow, TensorFlow, and reinforcement learning for scalable, personalized chatbot systems.

1.9 AI-Based College Inquiry Systems. Lalwani et al. [10]

This research work proposed a modular college inquiry chatbot that uses AIML, NLP-based keyword extraction, and semantic similarity measures to automate academic and administrative queries; it integrates user authentication, contextual intent detection, and adaptive knowledge retrieval for efficient domain-specific responses.

1.10 Immersive 3D Modeling Interfaces. Takala et al. [11]

This research work introduced an immersive 3D modeling framework that uses stereoscopic displays, head tracking, and motion-based controllers; their user study showed that spatial 3D interaction enhances immersive modeling.

1.11 AI-Driven Chatbot Architectures: A Systematic Review. Sharma et al. [12]

This research work presented a survey of chatbot frameworks with a focus on trends in rule-based, retrieval-based, and generative architectures, and demonstrated the shift toward hybrid NLP–ML models that enhance contextual understanding and scalability.

1.12 WebGL-Based 3D Interaction Using ThreeJS and Unity. Bheda et al. [13]

This research work demonstrated a WebGL pipeline using ThreeJS and Unity for interactive 3D environments in browsers, illustrating how lightweight rendering, optimized mesh pipelines, and WebGL shaders can enable real-time immersive experiences across devices.

1.13 Adaptive Chatbots with Dynamic Response Refinement. Reddy et al. [14]

This research work discussed adaptive chatbots that utilize NLP, reinforcement learning, and continuous user-feedback loops for self-improving conversational performance, including advancements in contextual embeddings, sentiment-aware processing, and autonomous dynamic response refinement.

1.14 Adaptive Emotional Intelligence in AI Chatbots. Rajuroy [15]

This research work explored how chatbots adaptively understand user affect, use NLP and reinforcement learning to adjust empathetic responses, and maintain interaction in culturally diverse markets.

1.15 AI as a Commoditized Technological Resource. Abonamah et al. [16]

This research work noted that the widespread availability of ML and deep learning tools is transforming AI into a commodity, reducing strategic differentiation, and suggesting the need for organizational frameworks that maintain competitive value.

1.16 Multilingual Chatbots for Indian Languages. Singh et al. [17]

This research work developed a multilingual chatbot using transformer-based models such as MuRIL BERT for fixed-response and context-based question answering, bypassing machine translation overhead and enhancing semantic accuracy for Indian languages.

1.17 Gamification for Engagement and User Retention. Koneti et al. [18]

This research work showed that gamified digital marketing systems using points, badges, leaderboards, and progression mechanics increase user engagement and emotional connection, while converting engagement into long-term retention.

1.18 WebGL for Browser-Based Graphics and Animation. Carson [19]

This research work described WebGL, a JavaScript API for rendering 2D/3D graphics in the browser, that is GPU-accelerated, and how it uses shaders and GPU-side computation to allow for real-time visualization and animation.

1.19 Progressive Web Apps and WebGL in Front-End Engineering. Venkataraajalu [20]

This research work discussed how PWAs, WebGL, and modern browser APIs are used together to create performant, interactive front-end applications that provide immersive, cross-platform web experiences using offline caching and GPU-accelerated rendering.

4 Objectives

The primary goals of this study are concentrated on improving human artificial intelligence psychological connection via intelligent multimodal design. The initiative is designed to produce a 3-D digital companion capable of showing sentiments spontaneously via coordinated visuals and realistic speech generation. It envisions incorporating a multilingual, sentimentally aware conversation framework that may adjust reactions centered on user voice and contextual sentiment. A further goal is to develop safe and continuous archives of consumer & pet interactions state using Firebase Firestore in order to guarantee continuity across devices. Furthermore, the study emphasizes on establishing gaming engagement measurements including gaming points, patterns, and psychological indicators to encourage long-term user involvement and emotional reinforcement. Besides fun, the structure intends to broaden such interpersonal interaction towards academic and medical purposes, concentrating on concern for others, emotional well-being, and accessibility.

5 Methodology and System Design

5.1 Architecture Design and Sequence Flow

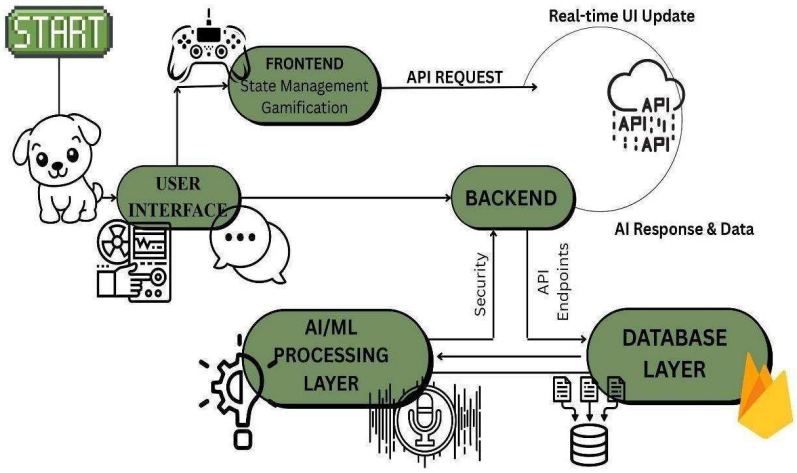

Fig. 1. Proposed system architecture and data flow.

Fig. 1. Demonstrates the structure of the digital companion system by illustrating how the consumer interface connects with the front-end for status handling and gaming, which routes API inquiries to the backend. The connections between the backend and the databases section for continuous information retention and the machine learning and artificial intelligence computation layer for intelligent reactions facilitate real-time revisions, reliable computation, and automatic cross-device synchronization.

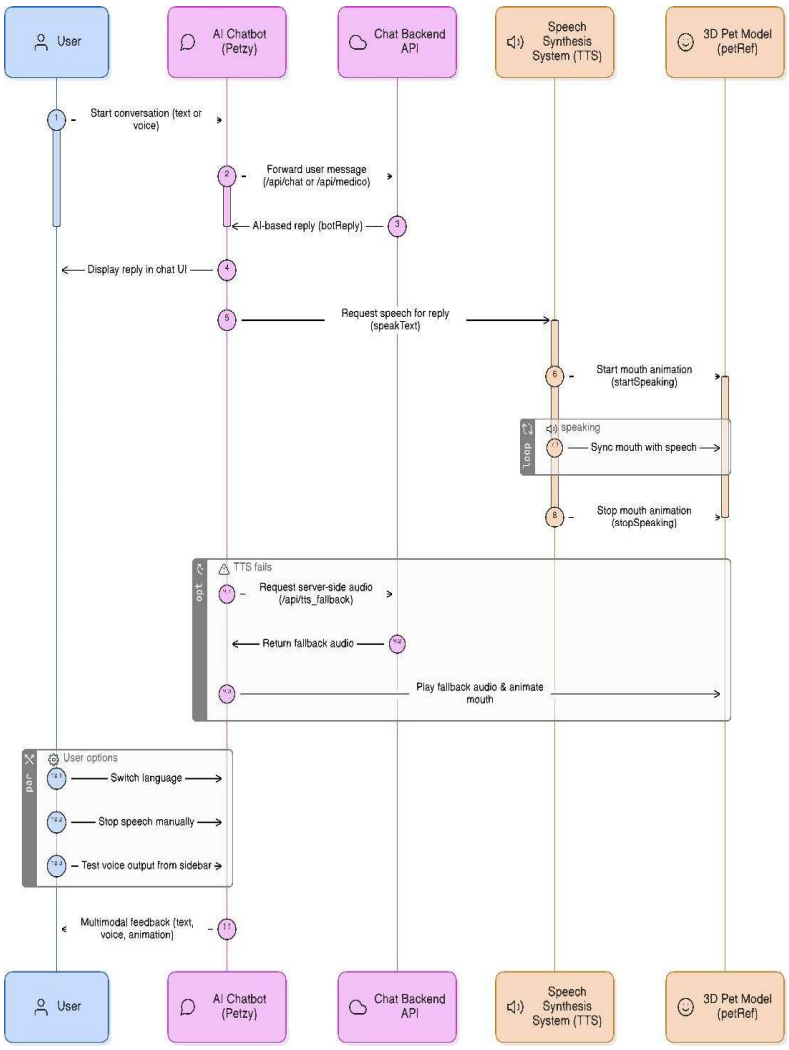

Fig. 2. Sequence Flow of AI-Powered Chatbot Interaction with Real-Time Mouth Animation.

Fig. 2. demonstrates the pattern of communications between the customer, the three-dimensional pet model, the messaging backend API, the AI chat agent (Petzy), and the voice synthesis system (TTS). It demonstrates how the chat bot interprets user information and communicates with the backend system to generate a perceptive response. The three-dimensional pet representation then coordinates its palate movements in actual time to correspond with the voice, whilst the TTS system converts the reaction into voice output. To guarantee constant interactions in the circumstance of a TTS connection failure, a substitute mechanism relies on server-generated audio. The effortless combination of AI conversations, voice transcription, and visual synchronization for an enjoyable digital pet adventure is illustrated by the picture, which also displays user choices like dialect alteration, and voice output examination.

5.2 Emotion Detection and Language Routing

Petzy's emotional recognition structure combines bilingual comprehension and psychological assessment to recognize both language and tone-related cues during interaction. The global process employs Unicode-driven language identification to autonomously recognize the content linguistic and connect it to the suitable natural language handling module. The emotion recognition is accomplished via a couple of rules-based and embedding-based models, allowing the system to distinguish psychological tone as positive,

neutral, or negative. These identified feelings directly impacts responses formation and graphical expressions, allowing the three-dimensional avatar to adapt its vocal tone and visual dynamics accordingly.

5.3 Speech and Interaction Loop

The actual time interactions loop in the Petzy framework integrates the user user interface, artificial intelligence predictive backend, and the three-dimensional rendering surroundings in a perpetual feedback cycle. Speech or written input from the individual is translated and investigated for sentimental and historical information beforehand being interpreted by the spoken AI module. The produced reaction is then presented both as written and coordinated speech with immediate mouth movements ensuring ideal alignment between speech and expression. This flawless blending of speech synthesis, psychological interpretation and three-dimensional modeling allows for realistic, psychologically aligned interactions between the customer and the digital companion.

5.4 Gamified State Management

Petzy encompasses gaming to increase user dedication and psychological connections via measurable interaction metrics. single player action, such as consuming food, engaging in conditioning, or transporting influences the pet's simulated state factors involving thirst, joyful energy, and affection. These efficient metrics are continuously updated and maintained in the cloud databases to preserve continuity across sessions. The pet's accomplishments and user participation are displayed through a real-time display that displays experience points (XP), activity periods, and psychological statistics using interactive data charts. This procedure reinforces continuous engagement and encourages positive social conditioning within the system.

5.5 3D Model and Customization

The visual illustration of the digital pet is created through high-quality three-dimensional modeling and animation. Blender was used to create multiple pet avatars which are portrayed within the React Three Fiber framework for interactive web deployment. Natural lighting, grain mapping, and dynamic shadows are handled by React Three producing flawless, lifelike movements at high frame rates. Users can customize their pets through color customization and habitat selection, improving emotional connection and a sense of belonging. This interactive aesthetically design helps contribute to Petzy's engaging quality and visual appeal.

5.6 Cross-Commodity Utility and Simulation Integration

The Petzy structure exhibits multipurpose utility, where the same emotional artificial intelligence construct supporting friendship is extended to health care simulations through the Medico Mode. This digital patient simulator employs Petzy's bilingual NLP, psychological analysis, and simulated feedback to simulate realistic patient interactions. It enables health sciences students to utilize medical dialogue and compassionate communication in a secure, controlled setting. The execution demonstrates Petzy's flexibility beyond entertainment, revealing its efficiency as a psychologically aware structure for structured medical training.

6 Results and Discussion

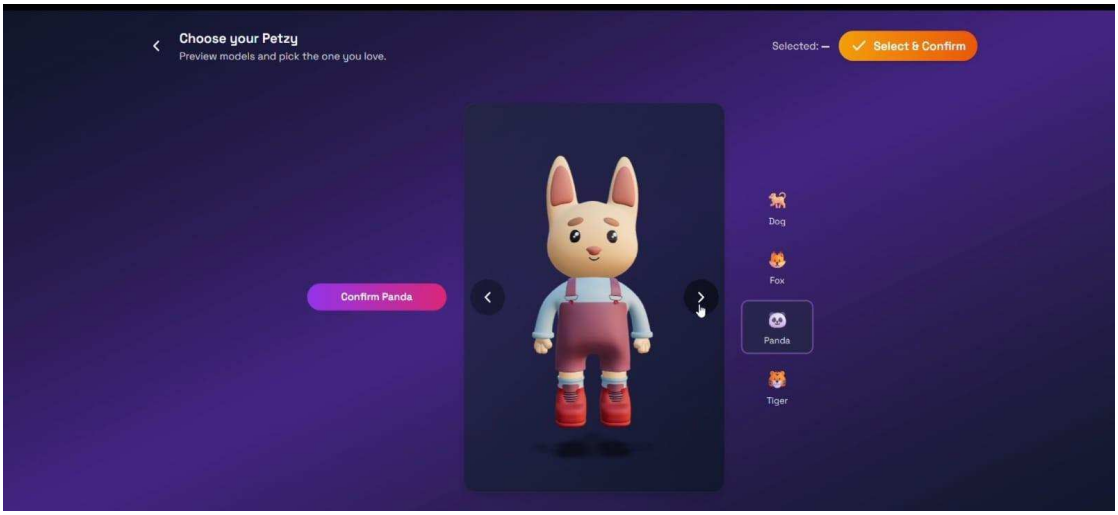

Fig. 3. Pet Model Selector.

This feature lets users see and select from four different 3D pet models as displayed in Fig. 3. Before moving on to customization or gameplay, users can choose their favourite model from an interactive preview of each model.

Fig. 4. Interactive Pet Dashboard and Key Functional Features of Petzy.

In addition to multilingual and contextually aware conversation, real-time psychological metrics, and behaviors like eating, playing, modifying, and teleporting, Petzy's interface combines Fun and Cognitive Zones. Petzy is an entertaining companion and a flexible tool for developing skills because of its Patient simulator and adaptive dialogue, which allow for realistic scenario-based learning.

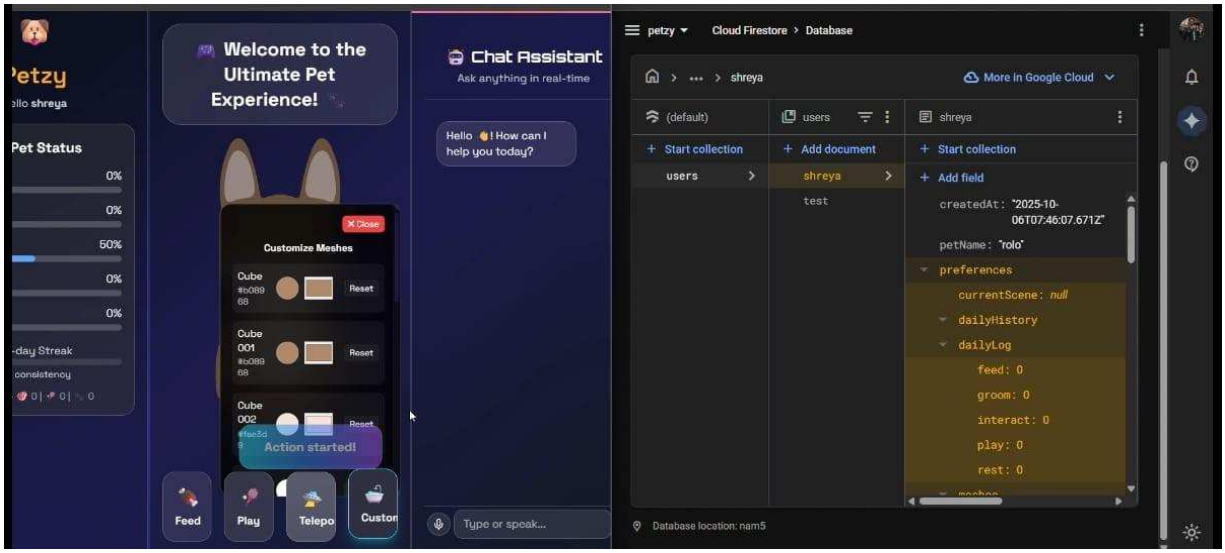

Fig. 5. No-Reload real-time data sync.

The no-reload instant data sync option makes sure that any storage changes such as pet behaviors or chat modifications are right away displayed on the user screen rather than requiring a page refresh. Real-time listeners from the Firestore are used to achieve this, offering a seamless and captivating user experience.

Fig. 6. User Dashboard.

The pet screen offers a real-time overview of the involvement, activity, and overall health of a digital pet. XP development and current action shifts are graphically analyzed using line and bar charts, and psychological indicators for starvation, happiness, and energy state are also shown. To offer users with an informative and clear summary of their pet's state, recent actions and metrics are revised constantly and a pie chart illustrates the breakdown of daily interactions.

Prototype verification showed that Petzy exhibited efficient and accurate real-time efficiency across its multimodal elements. The system preserved an ordinal response delay of roughly 1.3 seconds using the Groq prediction model and 1.8 seconds via the Ollama fallback, guaranteeing smooth conversational continuity. The speech identification precision reached approximately 93% for English and surpassed 85% for

significant Indic languages, indicating robust multilingual adaptability. Real-time communication through Firebase Firestore has enabled user user interface updates within 200 milliseconds without requiring page reloads, leading to seamless interaction. In addition, qualitative evaluations indicated strong emotional uniformity, with users assessing the AI friend as more compassionate and contextually adaptable than scripted chatbot systems, highlighting the efficiency of Petzy's emotionally adaptive structure.

Petzy illustrates that integrating psychologically intelligent AI with gamified communication promotes sustained player attachment and delivers plausible digital companionship. Its multifaceted architecture—comprising words, voice, and 3D interaction, crosses the distinction between practical and emotional computing. Key results include enhanced accessibility through bilingual support and automated script identification, compassionate response modulation facilitated by emotion and grammar analysis, and flexibility to healthcare fields through the "Medico Simulator" mode, that supports interpersonal skill training in virtual patient interactions. Current constraints involve reliance on web-based text-to-speech systems and the absence of constructive feedback processes for long-term acquisition of emotional patterns. Future improvements aim to employ reinforcement education to refine empathy adaptation, integrate motion-capturing data for better animation realism, and improve cross-domain training capabilities within academic and therapeutic contexts.

7 Conclusion and Future Scope

Petzy provides a novel structure for effective bilingual companionship that includes LLMs, 3D graphics, and gamification. The framework redefines virtual pet communication as an intellectually aware, context-adaptive experience. Further research will focus on virtual or augmented Reality incorporation, multi-agent cooperation for group learning surroundings, and psychological emotion perception (EEG or heart-rate inputs) to improve authenticity and therapeutic application. Through these enhancements, Petzy aims to serve as a flexible, inclusive and empathetic AI ecosystem.

References

1. Benner, Dennis & Schöbel, Sofia & Leimeister, Jan Marco. (2024). Engaging Minds – How Gamified Chatbots Can Support and Motivate Learners in Digital Education. 10.24251/HICSS.2024.008.
2. Hidayatulloh, Indra & Pambudi, Sigit & Surjono, Herman & Sukardiyono, Totok. (2021). Gamification on Chatbot-Based Learning Media: a Review and Challenges. Elinvo (Electronics, Informatics, and Vocational Education). 6. 71-80. 10.21831/elinvo.v6i1.43705.
3. Silva, Susana & De Cicco, Roberta & Levi, Maria & Hammerschmidt, Maik. (2023). Value Creation in Gamified Chatbot Interactions and Its Impact on Brand Engagement. 10.1007/978-3-031-25581-6_4.
4. Vicci, Heidrich. (2024). Emotional Intelligence in Artificial Intelligence: A Review and Evaluation Study. SSRN Electronic Journal. 10.2139/ssrn.4818285.
5. Ramsay, David B., Jiahui Jin, Pattie Maes and Roz Picard. "Virtual Pets and Virtual Selves as Exercise Motivation Tools." (2015).
6. Pearson, Andrew. (2020). Personalisation the artificial intelligence way. Journal of Digital & Social Media Marketing. 7. 10.69554/JJGR7331.

7. Basireddy, Maheswara. (2024). Implement Speech Recognition and Synthesis System Using Python. Journal of Mathematical & Computer Applications. 1-3. 10.47363/JMCA/2024(3)170.
8. Rastogi, Rohit & Anand, Tushar & Sharma, Shubham & Panwar, Sarthak. (2023). Emotion Detection via Voice and Speech Recognition. International Journal of Cyber Behavior, Psychology and Learning. 13. 1-24. 10.4018/IJCBPL.333473.
9. Mahmoud, Maha & Kumar, Ramakrishna. (2020). A Review on Chatbot Design and Implementation Techniques. International Journal of Engineering and Technology. 7. 2791.
10. Lalwani, Tarun & Bhalotia, Shashank & Pal, Ashish & Bisen, Shreya & Rathod, Vasundhara. (2018). Implementation of a Chat Bot System using AI and NLP. International Journal of Innovative Research in Computer Science & Technology. 6. 26-30. 10.21276/ijircst.2018.6.3.2.
11. T. M. Takala, M. Mäkäräinen and P. Hämäläinen, "Immersive 3D modeling with Blender and off-the-shelf hardware," 2013 IEEE Symposium on 3D User Interfaces (3DUI), Orlando, FL, USA, 2013, pp. 191-192, doi: 10.1109/3DUI.2013.6550243.
12. Wangsa, Ketmanto & Karim, Shakir & Gide, Ergun & Elkhodr, Mahmoud. (2024). A Systematic Review and Comprehensive Analysis of Pioneering AI Chatbot Models from Education to Healthcare: ChatGPT, Bard, Llama, Ernie and Grok. Future Internet. 16. 10.3390/fi16070219.
13. Giri, Prankush & Chandra, Sayanabha & Chakraborty, Subhrajyoti & Datta Choudhury, Panchali. (2023). Utilization of Three.JS and Unity to Make A Webgl API to Perform 3D Immersiveness In Web. American Journal of Electronics & Communication. 3. 20-23. 10.15864/ajec.3405.
14. Ram, Abhi & Salammagari, Abhi Ram Reddy & Srivastava, Gaurava. (2024). ADAPTIVE CHATBOTS: ENHANCING USER EXPERIENCE THROUGH INTERACTIVE LEARNING AND DYNAMIC RESPONSE REFINEMENT. 11.
15. Rajuroy, Adam & Isreal, Olatunji. (2023). Adaptive Emotional Intelligence in AI Chatbots: Enhancing User Engagement in Diverse Markets.
16. Abonamah AA, Tariq MU, Shilbayeh S. On the Commoditization of Artificial Intelligence. Front Psychol. 2021 Sep 30;12:696346. doi: 10.3389/fpsyg.2021.696346. PMID: 34659012; PMCID: PMC8514611.
17. Singh, Usneek & Vora, Nisarg & Lohia, Punit & Sharma, Yashvardhan & Bhatia, Ashutosh & Tiwari, Kamlesh. (2023). MULTILINGUAL CHATBOT FOR INDIAN LANGUAGES. 10.13140/RG.2.2.30993.20325.
18. Koneti, Chaitanya. (2025). “From Engagement to Retention: The Power of Gamification in Digital Marketing.”. Journal of Information Systems Engineering and Management. 10. 25-35. 10.52783/jisem.v10i4.8382.
19. Carson, Daniel. (2020). Research-Paper-WebGL. 10.13140/RG.2.2.18538.95683.
20. Ramadurai Venkataraajalu, Sivaramarajalu. (2024). The Role of Progressive Web Apps and WebGL in Modern Front-End Engineering. IJARCCE. 10. 10.17148/IJARCCE.2021.10830.