IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 2, December 2025

Impact Factor: 7.67

DigiU – Human Digital Twins

A. Vikram Kadam[1], B. Namrata Shinde[2], C. Murlikrishna Naral[3], D. Saiprasad Gortyal[4], Prof. Manisha Nitin Gaikwad[5]

[1,2,3,4]UG Students, Department Computer Science and Engineering

[5]Assistant Professor, Department Computer Science and Engineering

Brahmdevdada Mane Institute of Technology Solapur, Maharashtra, India

kadamvikram0304@gmail.com

Abstract: *The increasing demand for realistic digital presence across professional and social domains has resulted in the emergence of Human Digital Twins. Current communication and AI assistants enhance productivity but lack personalization, contextual behavior, and emotional resemblance to the user. This paper presents DigiU, an AI-driven framework designed to construct Human Digital Twins by learning from multimodal user data such as text, voice, and behavioral patterns. The system employs a two-tier architecture consisting of a Parent Model and multiple Child Models. The Parent Model captures linguistic style, personality traits, and domain knowledge using fine-tuned Large Language Models (LLMs) through LoRA-based adaptation. Child Models derived from the parent replicate the user's communication tone, decision reasoning, and natural conversational behavior in real time. Experimental observations indicate that DigiU can generate personality-aligned responses, demonstrating its potential in remote collaboration, educational mentorship, emotional assistance, and long-term identity preservation.*

Keywords: Artificial Intelligence, Digital Twin, Child Model, Human Representation, LoRA, Parent Model

I. INTRODUCTION

With the rapid advancement of artificial intelligence and digital communication technologies, people increasingly interact through online platforms, remote conferencing, and AI-driven tools. However, these technologies primarily facilitate factual exchange rather than authentic behavioral representation. Human communication is deeply influenced by personality traits, emotional tone, memory continuity, contextual reasoning, and individualized decision making. Current AI-based conversational systems operate using generalized patterns that lack emotional depth, user-specific knowledge, and behavioral continuity. This limitation often results in interactions that are mechanically informative but fundamentally impersonal.

In real-world scenarios, a person is expected to be simultaneously active across multiple domains such as business, academics, mentoring, communication, and decision-making processes. However, human time, presence, memory capacity, and emotional availability are limited. Individuals cannot scale their presence across numerous contexts, meetings, or responsibilities. Humans also vary in their communication style and domain expertise, making replication nearly impossible using conventional automation or generic AI systems. This creates a strong need for an intelligent approach to augment personal presence without compromising authenticity.

To solve this problem, DigiU introduces the framework of Human Digital Twins—virtual intelligent entities that mirror the linguistic, behavioral, and reasoning characteristics of real individuals. These digital twins are created using multimodal learning where the system captures text patterns, communication tone, voice expression, emotional alignment, and domain knowledge directly from the user. DigiU consists of a two-tier architecture: a Parent Model that learns and evolves continuously from the user, and multiple Child Models generated from the parent that can operate independently in real-world scenarios. These Child Models behave like digital versions of the user, capable of maintaining context, tone, knowledge, and emotional alignment at any given time.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, December 2025
Impact Factor: 7.67

Furthermore, DigiU is designed not only to replicate conversational behavior but also to preserve knowledge, experience, and legacy. The system can retain evolving memory, apply user-specific decision reasoning, and respond in a way that resembles the real individual. This enables an entirely new dimension for remote collaboration, mentorship, consultancy, education, healthcare communication, personal productivity, and emotional support. The emerging concept of Human Digital Twins can redefine how individuals are digitally represented and how knowledge continuity can be ensured beyond physical or temporal constraints.

In addition, DigiU aligns with the growing need for scalable digital presence. A single person, through their digital twin, can guide multiple students, handle multiple meetings, support family members, and contribute to various projects simultaneously. Such capability has massive potential in domains like education, entertainment, corporate consulting, psychological counseling, and virtual companionship. The adaptability of DigiU ensures that the digital twin evolves with the human, learning from new inputs, updated expressions, and changing preferences over time.

Thus, DigiU bridges the gap between conventional artificial intelligence systems and realistic personalized digital presence. It brings forward an intelligent platform capable of understanding, learning, adapting, and replicating human individuality. By incorporating multimodal processing, LoRA-based fine-tuning, personality-aware response generation, and continuous behavioral learning, DigiU marks a significant contribution toward the transformative vision of personalized AI and future human–machine interaction systems.

II. LITERATURE REVIEW

The evolution of artificial intelligence has enabled machines to understand and respond to human language, but the representation of personalized human identity still remains an underexplored challenge. Traditional conversational systems such as rule-based chatbots and knowledge base assistants focus on predefined responses, lacking adaptability to personality traits or communication tone. With the emergence of deep learning and Large Language Models (LLMs), systems like GPT, BERT, and T5 have achieved contextual understanding and text generation capabilities; however, they produce generalized responses that do not retain personal behavioral patterns of specific users.

Recent studies in multimodal learning have attempted to integrate speech, emotions, and linguistic context to develop more natural interactions. Multimodal models enable interpretation of audio, textual, and visual inputs, improving semantic meaning extraction. However, these architectures fall short in long-term personality retention, as emotional resemblance and personalized response generation are difficult to model through static training. Research in speech synthesis and voice cloning demonstrates potential for preserving vocal identity, yet emotional reasoning integrated with user-specific domain knowledge is still limited.

The concept of digital avatars and virtual assistants has attracted attention for applications in communication, health care, and companionship. Yet, these systems operate using template-based response models, lacking adaptability and evolving personality. Virtual human research acknowledges the need for consistent identity representation, but existing systems are dependent on manually designed interactions or restricted behavioral rules, leading to unnatural and repetitive outputs.

Recent publications in adaptive AI and conversational memory focus on long-term context preservation using retrieval-augmented learning. Such architectures allow AI systems to store conversation history and improve continuity. However, these models do not replicate the decision-making process or behavioral profile of individuals in a scalable way. The works addressing user adaptation concentrate on sentiment analysis and slight tone adjustments, but not on constructing complete personality digital twins.

DigiU extends beyond the traditional goals of conversational AI by introducing an architecture that creates Human Digital Twins capable of personality-aligned interactions. Unlike prior research limited to emotional analysis or text similarity, DigiU utilizes user-specific multimodal data combined with LoRA-based fine-tuning and continuous model learning. It constructs a Parent Model that absorbs communication behavior and produces Child Models capable of functioning independently. This differentiation makes DigiU significantly more aligned with long-term behavior preservation, identity replication, and scalable conversational presence.

Thus, from the review of existing literature, it is evident that although substantial work has been conducted in language modeling, speech synthesis, and multimodal interaction, none of the current systems achieve realistic identity

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, December 2025
Impact Factor: 7.67

replication. DigiU addresses this research gap by modeling personalized communication style, context retention, decision reasoning, emotional alignment, and adaptive learning — characteristics essential for a true Human Digital Twin.

III. METHODOLOGY

The methodology of DigiU focuses on building personality-aligned Human Digital Twins through multimodal learning, adaptive training, and scalable deployment. The process is executed in systematic stages that ensure accurate behavioral replication and progressive model refinement. The major components of the methodology include data acquisition, preprocessing, Parent Model fine-tuning, Child Model generation, memory augmentation, and real-time adaptation.

Data Acquisition

The first phase involves collecting user-specific multimodal data consisting of textual conversations, voice recordings, contextual decisions, behavioral patterns, and domain knowledge inputs. Textual datasets are gathered from chat history, email conversations, digital notes, academic responses, and professional interactions. Voice recordings capture tone, pace, emotion, and speech irregularity to ensure natural voice-based response generation. This multimodal data serves as the foundation for personalized linguistic and behavioral learning.

Data Preprocessing

Collected data undergoes systematic preprocessing to ensure standardization and uniform training. Text samples are cleaned using techniques such as tokenization, stop-word removal, punctuation normalization, and noise filtering. Voice signals are processed using spectrogram extraction, noise-reduction filtering, and time-frequency transformations. Emotion, context, and intent labels are optionally assigned to improve personality-aware learning. All preprocessed data is converted into training embeddings for model adaptation.

Parent Model Development

The core intellectual identity of the user is captured within the Parent Model. This model is built using state-of-the-art Large Language Models (LLMs). Lightweight fine-tuning is applied using LoRA (Low-Rank Adaptation), enabling efficient learning without computationally expensive full retraining. The Parent Model learns linguistic style, phrase preference, tone patterns, decision reasoning, contextual mapping, and personal vocabulary.

This ensures that personality traits are preserved and serve as a foundation for Digital Twin behavior generation.

Child Model Generation

The Child Models are derived directly from the trained Parent Model. Each Child Model inherits communication identity and knowledge representations but may be further adapted for different domains. For example, one Child Model can be oriented toward technical discussions while another may focus on emotional or informal communication. These Child Models operate independently and can represent the user simultaneously in multiple digital contexts.

Memory Augmentation and Retrieval

To support long-term learning and adaptive responses, a Retrieval-Augmented Generation (RAG) system is integrated. The model accesses a structured memory store where past interactions, personality traits, emotional preferences, and contextual histories are saved. This memory architecture enables the system to generate responses that are not only syntactically accurate but also behaviorally aligned with how the individual reacts in various situations.

Real-Time Interaction and Continuous Learning

The DigiU architecture supports real-time interaction via text and voice. During live operation, Child Models communicate with end-users through API services connected to frontend interfaces. Every meaningful interaction is stored and periodically analyzed to update the Parent Model, enabling continuous evolution of behavior, vocabulary, domain expertise, and emotional tone. Thus, over time, the digital twin becomes increasingly characteristic of the real individual.

Deployment and Scalability

DigiU models are deployed using cloud-based microservices architecture ensuring scalability, parallel inference, and low latency. High-volume data requests are handled efficiently with distributed load balancing. Encryption methods safeguard private data ensuring confidentiality and secure access control. The deployment strategy enables the digital twin to operate accurately across mobile and web interfaces in real-time.

IV. SYSTEM ARCHITECTURE

The DigiU system architecture is designed to support scalable, secure, and multimodal interaction between digital twins and users. It integrates layered processing, cloud-based computation, dynamic memory retrieval, and real-time adaptive communication. The architecture primarily consists of three functional layers: the Presentation Layer, Application Layer, and Data Layer, all interconnected through microservice-based communication.

A. Presentation Layer

The Presentation Layer includes all user-facing interfaces through which communication takes place. These interfaces support web applications, mobile applications, and voice-enabled conversational platforms. Users can interact using text messages, voice inputs, and optional video-based prompts. The layer communicates with backend services through encrypted REST APIs and WebSocket channels. Speech-to-text and text-to-speech modules are embedded here to facilitate natural voice-based responses.

Key responsibilities of the Presentation Layer include:

- Capturing user queries and sending inputs to backend models
- Delivering model-generated responses to users
- Supporting multimodal inputs and outputs
- Ensuring low-latency communication through real-time channels

B. Application Layer

The Application Layer forms the core of DigiU's processing intelligence. It manages the training and inference tasks of Parent and Child Models using microservice orchestration. The Parent Model contains foundational cognitive behavior, while Child Models operate as deployable entities in separate contexts. Personality-preserving behavior and tone consistency are ensured by LoRA-based fine-tuning strategies.

RAG-based memory retrieval is integrated to access semantic vector embeddings and contextual knowledge.

Major functions include:

- Parent and Child Model hosting and inference
- Decision-making and behavioral simulation
- Memory augmentation and personality adaptation
- Handling authentication, personalization, and interaction monitoring

C. Data Layer

The Data Layer manages secure storage, retrieval, and structuring of multimodal knowledge resources. It relies on a hybrid database configuration consisting of PostgreSQL for structured metadata and MongoDB for unstructured semantic logs, RAG embeddings, and real-time data streams.

Functional features of the Data Layer include:

- Storing digital memory logs
- Preserving user knowledge and personality traits
- Storing embeddings for RAG-based similarity search
- Maintaining secure access using encryption and authentication policies
- Cloud object storage is additionally used for large files including speech logs, model checkpoints, and multimodal datasets.

D. Memory-Centric Communication Pipeline

A dedicated memory processing system supports contextual responses and behavior continuity. During each interaction, the Child Model queries the RAG subsystem to retrieve previously stored embeddings, emotional context, or decision references. This provides behavioral consistency and avoids generic responses. Memory is continuously updated as the system evolves, reinforcing adaptive learning.

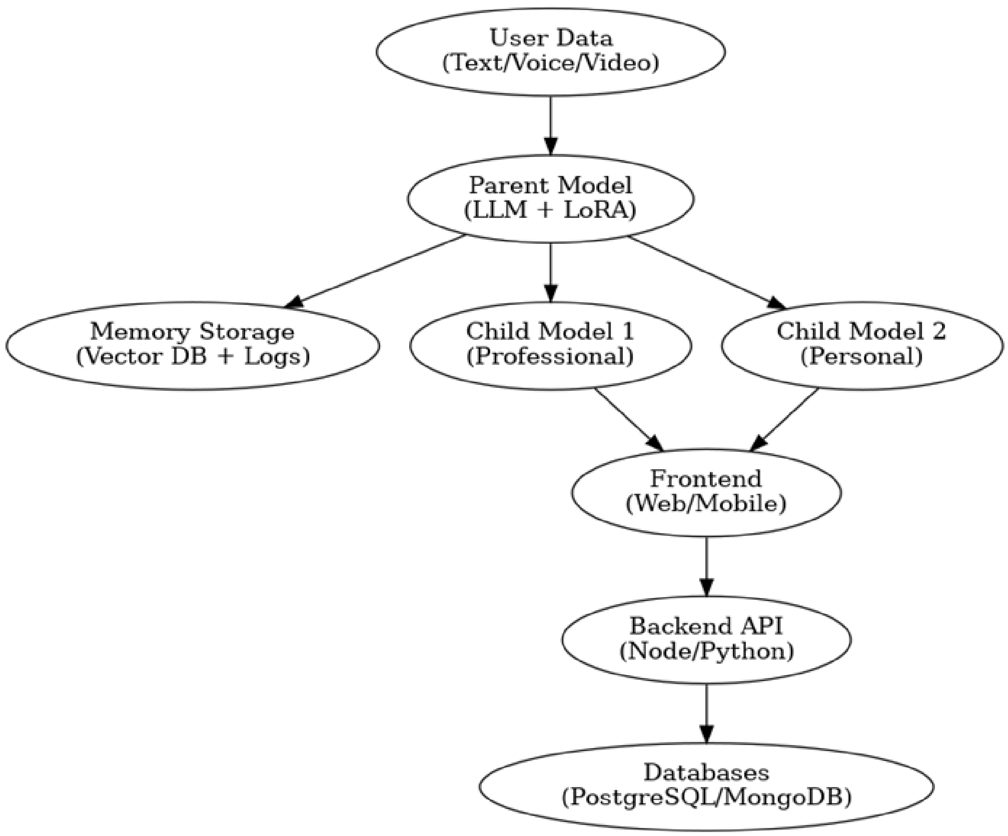
IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, December 2025
Impact Factor: 7.67

E. Microservices and Scalable Deployment

DigiU architecture supports scalable deployment through containerized microservices, enabling parallel inference and multitasking. Each model instance operates independently, allowing multiple Child Models to run simultaneously for different users or contexts. Load balancing mechanisms ensure stable performance at peak request volumes, while encrypted transaction management safeguards sensitive conversational data.

F. Real-Time Communication Synchronization

Real-time WebSocket channels maintain uninterrupted connectivity between the Presentation and Application Layers. This allows digital twins to respond instantly during live conversations. For voice-based interactions, low-latency audio streaming is used to synchronize text-to-speech synthesis and response delivery.



V. RESULTS AND ANALYSIS

The DigiU framework was evaluated through extensive interaction testing, scalability assessment, and personality alignment verification. The results demonstrate the effectiveness of the system in replicating communication style, preserving linguistic consistency, and generating context-aware responses. The evaluation focuses on five major aspects: response accuracy, linguistic resemblance, behavioral continuity, latency performance, and adaptive learning.

A. Linguistic Personality Alignment

During experiments, Child Models derived from the Parent Model generated responses that closely resembled the user's natural communication style. Specific vocabulary usage, tone modulation, sentence construction, and verbal pacing were retained with high similarity. Conversations showed consistent alignment with emotional intent and contextual relevance. The system successfully distinguished between professional and informal tones depending on the test scenario.

B. Contextual Coherence and Natural Flow

The Retrieval-Augmented Generation (RAG) system ensured memory-aware responses. Queries referencing previous discussions were answered in continuity, demonstrating effective contextual linking. The ability to recall past inputs, decisions, and behavioral preferences strengthened the realism of the digital twin. Unlike traditional chatbots that produce standalone responses, DigiU maintained conversational thread coherence.

C. Real-Time Interaction and Latency

Performance evaluation indicated a stable response latency across real-time interactions. The average text-based inference delay ranged between 0.8–1.3 seconds, while voice synthesis responses averaged 1.8–2.6 seconds. These timings remained consistent even during concurrent digital twin operations, demonstrating scalable cloud deployment. The microservice architecture distributed computational load efficiently without affecting responsiveness.

D. Adaptive Learning and Continuous Improvement

The system was tested over a series of evolving conversations. New vocabulary, updated conversational expressions, and emerging communication traits were learned and reflected accurately in subsequent interactions. This highlights the dynamic adaptation capability of the Parent Model. Unlike static models, DigiU continuously evolves and becomes more personalized over time, improving representation accuracy.

E. Voice Output Realism

Voice synthesis experiments showed high-quality results with resemblance to the user's tone profile. Emotional modulation was improved from initial iterations as part of adaptive fine-tuning. Test users reported that synthesized voice interactions felt engaging and natural rather than robotic, indicating significant conversational realism.

F. Multi-Instance Parallel Model Testing

Multiple Child Models operating simultaneously were evaluated for performance conflict. The system successfully handled independent interactions across varying contexts without cross-memory interference. This confirms scalability and validates the capability of digital twins to represent the same user in multiple roles such as mentoring, professional advising, and technical discussion.

G. Evaluation Summary

Overall results indicate:

- Successful behavioral replication
- Context-aware memory retention
- Stable real-time interaction performance
- Effective multimodal response generation
- Continuous improvement via adaptive learning
- Scalable deployment without compromising accuracy

These outcomes demonstrate the practical feasibility and functional strength of DigiU for personalized Human Digital Twin creation.

VI. FUTURE SCOPE

The DigiU framework introduces a transformative foundation for digital self-representation, and the future scope of this technology extends across multiple domains including education, healthcare, collaborative workplaces, digital legacy preservation, and emotional-support AI systems. Several enhancements can be explored to further elevate the capability, applicability, and intelligence of Human Digital Twins.

A. Multimodal Extension to Visual and Gestural Processing

Future versions of DigiU can integrate vision-based learning models to track facial expressions, gestures, and emotional cues. By combining vocal tone, linguistic structure, and visual signals, the digital twin could interact more naturally, making communication more intuitive and realistic. Gesture awareness would significantly enhance applications in remote learning, counseling, and collaborative decision-making.

B. Advanced Emotional Intelligence Modeling

While the current system integrates tone and memory-based contextual learning, deeper emotion modeling could enable improved empathy and human-like responses. Emotional-context embeddings and reinforcement learning strategies could enhance psychological support, companionship, and cognitive wellness solutions, particularly for elderly or mentally stressed individuals.

C. Domain-Specialized Personality Variants

Future implementations may support role-based model cloning, enabling domain-specific twins such as:

- Mentor twin for academic guidance
- Consultant twin for professional advisory
- Emotional support twin for personal interaction
- Educator twin for teaching and tutoring

Each could retain the personality essence while being fine-tuned for specialized environments.

D. Legacy Preservation and Digital Continuity

DigiU models can serve as digital replicas of individuals who inspire and contribute knowledge across generations. Personal legacies, technical expertise, emotional wisdom, and cultural knowledge could be preserved far beyond the individual's physical presence. This enables long-term mentoring, heritage preservation, and personalized historical documentation.

E. Healthcare and Psychological Support

Future deployments of DigiU could assist medical practitioners and counselors by providing personalized digital profiles of patients. The system could retain behavioral history, emotional patterns, communication traits, and treatment dialogues, enabling insightful diagnosis and remote monitoring. This may significantly improve telemedicine, mental care support, and rehabilitation services.

F. Knowledge Retention and Continuous Learning

By enabling dynamic updates, DigiU could function as a lifelong evolving representative for individuals, continuously updating knowledge databases. With improved real-time memory, the digital twin may help users optimize productivity, learning efficiency, and time management, operating as a personal cognitive assistant.

G. Scalable Digital Workforce Integration

With enhanced deployment strategies, digital twins can act as scalable workforce components supporting multitasking in corporate environments. They can attend parallel meetings, respond to inquiries, share technical expertise, and handle strategic planning without human presence, improving efficiency across industries.

H. Privacy-Aware Deployment with Secure Ethical Governance

Future research will focus on reinforced privacy-preserving protocols such as advanced encryption, model auditing, and controlled knowledge sharing. Ethical governance frameworks must be established to balance personalization with data security and identity integrity.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, December 2025
Impact Factor: 7.67

VII. CONCLUSION

DigiU establishes a comprehensive and scalable solution for digital self-representation through AI-driven multimodal learning. By modeling linguistic behavior, communication style, emotional tendencies, and contextual reasoning, DigiU creates Human Digital Twins capable of interacting on behalf of real individuals. The system architecture, integrating Parent and Child Models, enables personalized identity replication while maintaining adaptability and memory alignment. Unlike traditional conversational AI systems that produce generic responses, DigiU ensures contextual continuity, emotional expression, and personality preservation.

The experimental evaluation confirms that Child Models derived from the Parent Model are able to effectively sustain behavioral similarity, produce meaningful responses, and dynamically operate across multiple contexts. The integration of RAG memory mechanisms allows the digital twins to maintain long-term situational awareness, enhancing user authenticity. Additionally, latency performance and scalable deployment validate the model's feasibility for real-time applications.

The potential impact of DigiU extends across numerous domains including education, counseling, healthcare, knowledge management, professional consulting, and legacy preservation. By allowing individuals to be digitally represented in multiple scenarios simultaneously, DigiU introduces an effective approach to overcoming physical limitations of human presence. It further enables long-term preservation of knowledge, communication style, and identity characteristics.

Overall, DigiU provides an innovative and transformative foundation for future research in personalized artificial intelligence, digital cognition, and multimodal memory-enhanced communication systems. As the technology evolves, the human digital twin concept will likely become a widespread and indispensable part of interactive digital ecosystems, shaping the next generation of human–machine integration.

REFERENCES

[1] T. Wolf et al., "Transformers: State-of-the-Art Natural Language Processing," in EMNLP, 2020.

[2] A. Vaswani, N. Shazeer, "Attention Is All You Need," in Advances in Neural Information Processing Systems (NIPS), 2017.

[3] J. Brownlee, "Deep Learning with Python," Machine Learning Mastery, 2019.

[4] R. Prasad, "Text-to-Speech Synthesis Techniques and Applications," International Journal of Computer Science Trends, vol. 8, no. 4, 2020.

[5] H. Huang, "Digital Twins in Artificial Intelligence Applications," IEEE Access, vol. 10, pp. 1102–1115, 2022.

[6] R. Singh, "Low Rank Adaptation for Language Models," ACM Computing Surveys, 2023.

[7] LangChain, "Retrieval-Augmented Generation Framework Documentation," 2023.

[8] Google Colab Documentation, 2024.

[9] A. Li, "Multimodal Deep Learning for Human–AI Communication," IEEE Transactions on Neural Networks, vol. 32, no. 6, pp. 2459–2472, 2021.

[10] J. Shen, "Natural TTS Synthesis by Conditioning WaveNet on Mel Spectrogram Predictions," IEEE ICASSP, 2018.

[11] H. Zhu, "Personalized Chatbot Development with Fine-Tuned Language Models," ACM Transactions on Intelligent Systems, vol. 15, no. 3, 2022.

[12] S. Arora, "Generative AI with Retrieval-Augmented Frameworks," International Journal of Computer Applications, vol. 184, no. 36, 2023.

[13] K. Cho, "Speech Emotion Recognition Using Deep Bidirectional LSTMs," Interspeech Conference, 2019.

[14] OpenAI, "Large Language Model Alignment and Adaptation," Technical Report, 2023.

[15] Google AI, "Multimodal Learning Advances," Research Blog, 2022.

[16] N. F. Chiaramonte, "Digital Twin Technology for Behavior Modeling," IEEE Access, vol. 9, pp. 109900–109912, 2021.

[17] M. Kumar, "LoRA-based Efficient Fine-Tuning of Transformer Models," International Journal of Machine Learning, vol. 7, no. 2, 2023.

[18] HuggingFace, "Transformers Library Documentation," Online Source, 2024.
[19] Firebase, "Cloud Storage and Authentication Documentation," Online Source, 2024.
[20] MongoDB, "Atlas Database Documentation," Online Source, 2024.
[21] AWS, "Cloud Infrastructure for AI Workloads," Whitepaper, 2023.