Prototyping a Personalized Chatbot Serving as a Game Facilitator for Older Adults

Y. C. Chiang, B. Q. Zhang, J. M. Cheng, Y. L. Hsu. *Gerontechnology* 25(s)

Purpose Serious games and exergames have been shown to improve cognitive and physical functions in older adults, helping maintain memory, attention, reaction time, and motor abilities [1,2]. Many successful game-based interventions rely on a capable facilitator who can provide timely guidance, encouragement, and individualized support to sustain engagement. However, most digital game applications rely on fixed, prerecorded instructions that cannot replicate the adaptive role of a human facilitator. This study aims to transform a traditional game App into a personalized game chatbot that serves as a simulated game facilitator for older adults. The system uses generative AI to create facilitator-style guidance, produce context-aware feedback, and deliver personalized voice output that reflects the facilitator's speaking characteristics. **Method** The system consists of three components.(1) Personalized voice model creation: The human facilitator records at least 10 seconds of speech. The audio is transmitted to the ElevenLabs AI voice platform (https://elevenlabs.io/) to establish a personalized voice model, which returns a dedicated Voice ID. Whisper API transcribes the facilitator's sample audio and extracts timestamps. Speaking rhythm and pauses are calculated and converted into a personalized speech-rate multiplier between 0.7 and 1.2, which is later used in speech synthesis. (2) Game chatbot interaction: As shown in Figure 1, the game App connects to the GPT-4o-mini Realtime model through WebSocket. Based on the game state, the model generates facilitator-style instructions and feedback, while handling context and situation modelling to produce relevant and adaptive speech output. The text and the speech-rate parameter are sent to ElevenLabs for speech synthesis using the personalized voice model and Voice ID.(3) Low-latency sentence-level synthesis: A sentence-level Text-to-Speech (TTS) mechanism synthesizes the next sentence while the current one is being played. This parallel process helps maintain smooth and responsive interactions during gameplay. A quiz game was chosen to build the first prototype to test this concept because it requires extensive verbal interaction between the player and the facilitator. We conducted three test sessions with three facilitators, covering the entire process from voice recording to actual gameplay.

Results Figure 2 shows the 10-question quiz game used to evaluate the prototype system. Across the three test sessions, the chatbot successfully presented questions, responded to player selections, and generated performance summaries using personalized voice models and individualized speech rates. Early user responses indicated that the personalized speaking rhythm improved naturalness and that the cloned voice increased familiarity. Participants also reported that the chatbot resembled the support provided by a human facilitator, which contributed to a more engaging gameplay experience for older adults. Formal quantitative data collection has not yet been conducted. **Discussion** Effective facilitation requires appropriate timing, rhythm, tone, and adaptability. This study combines voice cloning, personalized speaking rhythm, and dynamically generated feedback to enhance the naturalness of game facilitation. By integrating these elements, the system produces natural, facilitator-like speech that more closely aligns with how a human facilitator interacts with older adults. This approach supports engagement in quiz-based serious games and enhances the overall user experience. Future work will involve rigorous field trials in daycare centers with caregivers and older adults to evaluate effectiveness, user acceptance, and long-term applicability in real-world settings. Future work may also integrate emotional tone to express encouragement or gentle correction at suitable moments. In addition, the personalized facilitator-style chatbot may be extended to cognitive training, social interaction support, and technology-assisted care scenarios for older adults.

References

1. Dwi Pratiwi, G., Hadiyani, W., Lucya, V., & Gumelar, Y. (2024). Exergame: Exercise With Matching Memory Game On Cognitive Function And Physical Fitness Of The Elderly.
2. Wang, S. M., Kang, D. W., Um, Y. H., Kim, S., Lee, S., Lee, C. U., & Lim, H. K. (2024). Effects of Serious Games in Older Adults With Mild Cognitive Impairment. *Psychiatry Investigation*, *21*(5), 449.
3. Yang, Z., Hong, J., Pang, Y., Zhou, J., & Zhu, Z. (2025). ChatWise: A Strategy-Guided Chatbot for Enhancing Cognitive Support in Older Adults. *arXiv preprint arXiv:2503.05740*.

Keywords: Personalized Voice Chatbot, Serious Games, Voice Cloning, Generative AI

Affiliation: Gerontechnology Research Center, Yuan Ze University, Taiwan
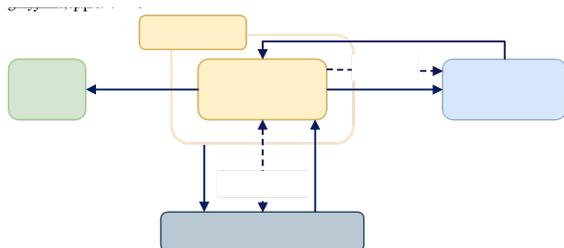**Email:** mehsu@saturn.yzu.edu.tww

Figure 1. Voice guidance flow

Figure 2. Quiz game interface