Demo: An Agentic Multi-Persona Generative AI System for Mental Health Companionship

Yuekai Wang
Rancho Bernardo High School
San Diego, CA 92128
kevinyuekai@gmail.com

Abstract

Loneliness among older adults is a significant public health concern associated with depression, cognitive decline, dementia, and premature mortality. Existing generative AI chatbots for mental health support often rely on a single persona, have limited memory, depend solely on pre-trained knowledge, and lack design considerations for older users. This demo presents HeartBond, an agentic AI system that supports multi-persona interaction with persistent memory and tool augmentation. HeartBond enables users to configure and interact with multiple custom AI personas that emulate the voices, personalities, and interests of people they care about. The goal of this demo is to showcase the working system, highlight how personalized generative AI can foster more natural and emotionally resonant companionship, and spark discussions on the potentials and challenges of applying generative AI to companionship and mental health.

1 Introduction

Social isolation and loneliness are serious public health risks that affect a large portion of the older adult population [17]. One in three adults aged 50 to 80 report persistent feelings of loneliness or isolation [5], a condition strongly associated with increased risk of depression, cognitive decline, dementia, and even early death [6]. Financial restrictions and geographic barriers make the issue worse. Many older adults struggle to afford the cost of hiring a full-time caregiver or moving into a senior home [7]. At the same time, a large proportion wishes to remain in their own house and familiar environments where they have lived for many years [9]. However, their adult children may live far away or face high demands from work and their own families, leaving seniors with limited day-to-day support and companionship.

Existing generative AI–based applications of large language models (LLMs) in mental health care can be categorized into six main types based on functionality: clinical assistant, counseling, therapy, emotional support, positive psychology intervention, and education [10]. Among these, emotional support applications aim to provide empathetic responses and psychological support in various contexts, such as alleviating loneliness or reducing suicide risks. Despite their strengths, current systems exhibit several notable limitations. Many rely on a single LLM persona with a uniform conversational tone, which constrains the personalization of interactions. Others suffer from limited context retention and operate with only session-bound memory, reducing conversational continuity. Some depend solely on pre-trained model knowledge, which limits their ability to provide contextually relevant and up-to-date assistance. Furthermore, most existing designs are not tailored to the needs of older adults, particularly those who may have limited technological proficiency.

To address these limitations, we introduce HeartBond, a multi-persona, voice-first generative AI system to provide immersive emotional companionship. The system uses a modular orchestrator-

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: The Second Workshop on GenAI for Health: Potential, Trust, and Policy Compliance.

worker agentic workflow, in which a central orchestrator coordinates multiple worker agents. Each worker-agent represents a persona and can be customized to replicate the voice, personality, and interests of a specific individual meaningful to the user, such as family members or close friends. The architecture is designed for flexibility, which allows new personas to be added without retraining. Additionally, the system is optimized for voice-first interaction but also supports text input to accommodate user preferences. Chat memory is shared among all agents, enabling them to recall past conversations and maintain continuity while gradually shaping more personalized interactions. In addition, the system uses real-time tool augmentation to retrieve external information via APIs or custom sources, ensuring responses that are accurate, contextually relevant, and up to date.

By integrating multi-persona interaction, persistent memory, voice-first communication, agentic orchestration, and dynamic tool augmentation, HeartBond advances beyond existing mental health chatbots. The goal of this demo is to showcase our working system, highlight its technical design, and spark discussion on its usability, safety, and ethical considerations.

2 Related Work

2.1 Generative AI-Based Conversational Agents for Mental Health

Generative AI mental health chatbots form a subset of conversational agents that employ LLMs to simulate dialogue. By moving beyond the constraints of traditional rule-based, they enable more natural, adaptive, and context-aware interactions. Serena [2] is a 2.7B-parameter sequence-to-sequence transformer fine-tuned on therapy transcripts, incorporating post-processing to enhance coherence and empathy. Similarly, Psy-LLM [12] integrates pre-trained LLMs with psychologist Q&A archives and mental health literature, demonstrating the growing sophistication of these tools. For peer support, Hailey [19] adopts an AI-in-the-loop approach to provide real-time empathic feedback to peer supporters in online mental health chats, Marmol-Romero et al. [16] examines a chatbot's engagement with Spanish-speaking teenagers on mental health topics, and Replika [15] introduces a GPT-powered chatbot that enables open-ended, empathetic conversations for loneliness and suicide mitigation support. However, despite their effectiveness for engagement, these systems remain limited by their focus on single-persona, primarily text-based interactions, and lack the memory and multimodal capabilities featured in modern generative AI technologies.

2.2 Multi-Agent and Multi-Persona Dialogue Architectures

Early persona-based system largely focused on static, single-character systems. Later, datasets such as PersonaChat [22] introduced predefined text-based persona profiles to promote consistent and engaging responses, but these profiles lacked the adaptability required for long-term evolving relationships. Multi-Character Chat [20] extends the concept by enabling multiple fictional characters to converse within a shared space, typically for narrative purposes. While this demonstrates the feasibility of role-switching, it generally lacks shared memory across characters and offers minimal real-time adaptability, limiting continuity and coherence over time.

Recent systems have explored explicit multi-agent orchestration to enhance personalization and realism in mental health contexts. MentalAgora [13] employs multiple agents, each embodying a distinct therapeutic approach, to debate and collaboratively construct a persona aligned with user preferences and needs, and MIND [3] assigns role-specific LLM agents within a structured healing framework to generate emotionally resonant dialogues. However, both systems are primarily designed for structured text-based therapeutic exchanges and do not incorporate persistent memory, multimodal voice interaction, or elderly-focused accessibility features that are central to HeartBond's emotionally supportive design.

Another broader set of conversational frameworks leverage agentic AI for multi-agent coordination even when persona-driven interaction is not the main objective. MAGI [1] uses specialized agents to manage the flow of interviews, generate questions, validate judgments, and produce diagnostic results. The AI-Assisted Multi-Agent Dual Dialogue System [11] similarly orchestrates clinician-facing agents to summarize dialogues, recommend exercises, analyze conversational themes, and suggest empathetic responses, with a human-in-the-loop for safety. While these architectures demonstrate scalability and task specialization, their focus on supporting the clinician, rather than supporting the user, limits their capacity for sustained emotional engagement. They also lack a persistent memory

and voice-first interaction, both of which are essential for fostering trust and long-term connections in emotionally supportive contexts, especially for seniors.

2.3 Generative AI in Elder Care

Recent research highlights the value of voice-based conversational agents in supporting the well-being and engagement of older adults. Studies show that voice-based conversational agents can integrate into daily routines, fostering emotional connection, a sense of competence, and positive experiences beyond basic functionality [18]. FitChat [21], a voice-based AI chatbot for promoting physical activity, demonstrated that older adults preferred voice interaction over text, finding it more motivating and easier to use. Similarly, research on voice assistants shows that they serve both assistive and social roles, offering hands-free convenience and perceived social support that increase adoption and satisfaction among seniors [8]. These findings demonstrate that voice-first interaction significantly enhances usability, accessibility, and emotional support for older adults, which is one of our core principles of HeartBond's design.

3 System Overview

3.1 System Design

HeartBond is developed using the OpenAI Agents SDK framework and coded in Python, with the UI mockup created using Gradio.

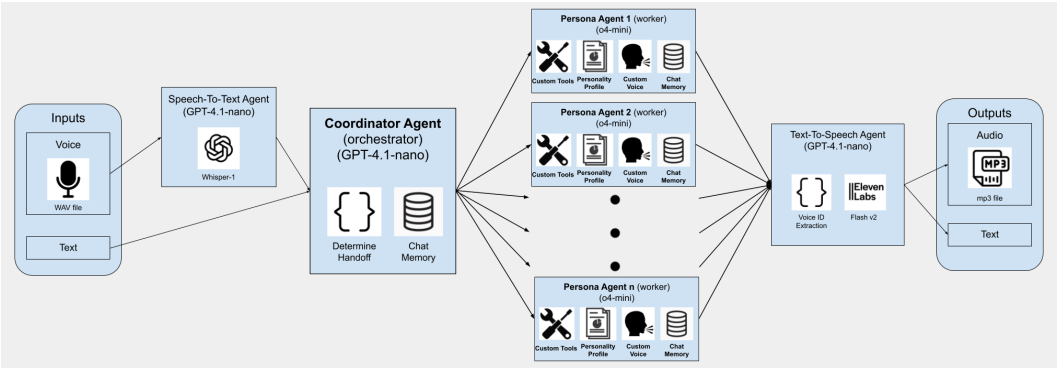

Figure 1: Workflow diagram of HeartBond's system architecture.

Each part of the workflow in Figure 1 is as follows:

Inputs: Users can interact via voice or text in the UI. Voice is saved as a WAV file.

Speech-to-Text Agent: If voice input is used, this agent transcribes the WAV file into text using OpenAI Whisper-1.

Coordinator Agent: Serves as the central orchestrator. It analyzes the user's input, decides which persona agent should respond, and performs the handoff to that agent. It also updates the UI to reflect the persona it is interacting with. Chat memory is maintained with OpenAI Sessions, allowing context-sensitive decisions across multiple turns.

Persona Agents: Configurable worker agents that emulate people chosen by the user, such as a spouse, friend, or family member. Each persona includes:

- **Personality Profile:** The persona's personality is described in the agent's prompt, capturing traits such as sociability, sense of humor, and conversational style. The persona's hobbies or interests are also included in the prompt, which correspond to the custom tools. Each persona is assigned a unique Voice ID which is trained in the custom voice module and passed to subsequent agents via prompt engineering.
- **Custom Tools:** Custom tools can make API calls to external sources that relate to the person's hobbies, such as a recipe API for someone who loves to cook. Instead of an API,

custom tools can also be personalized to the user, such as retrieving from a database of that person's personal recipes.

- **Custom Voice:** The voice is trained using ElevenLabs and requires a minimum of ten seconds of recorded speech from the person, provided by the user. After training, the persona is assigned a specific Voice ID.
- **Chat Memory:** All persona agents share access to the same conversation history, which allows them to maintain continuity and context in multi-turn dialogue.

Text-to-Speech Agent: After a persona generates a response, this agent retrieves the associated Voice ID and calls the ElevenLabs API to generate mp3 audio that matches the persona's voice.

Outputs: The persona's response audio plays in the UI alongside the displayed text output.

3.2 Use Case Demo

We demonstrate HeartBond through a use case for *Susan*, a lonely grandma who wishes to connect with three people: her husband *George* who is deceased, her close friend *Maria* who moved away, and her grandson *Alex* who is at college, shown in Figure 2. For demo purposes in this paper, interactions are shown in text dialogues, but more details and a video with audio can be found in Appendix A.

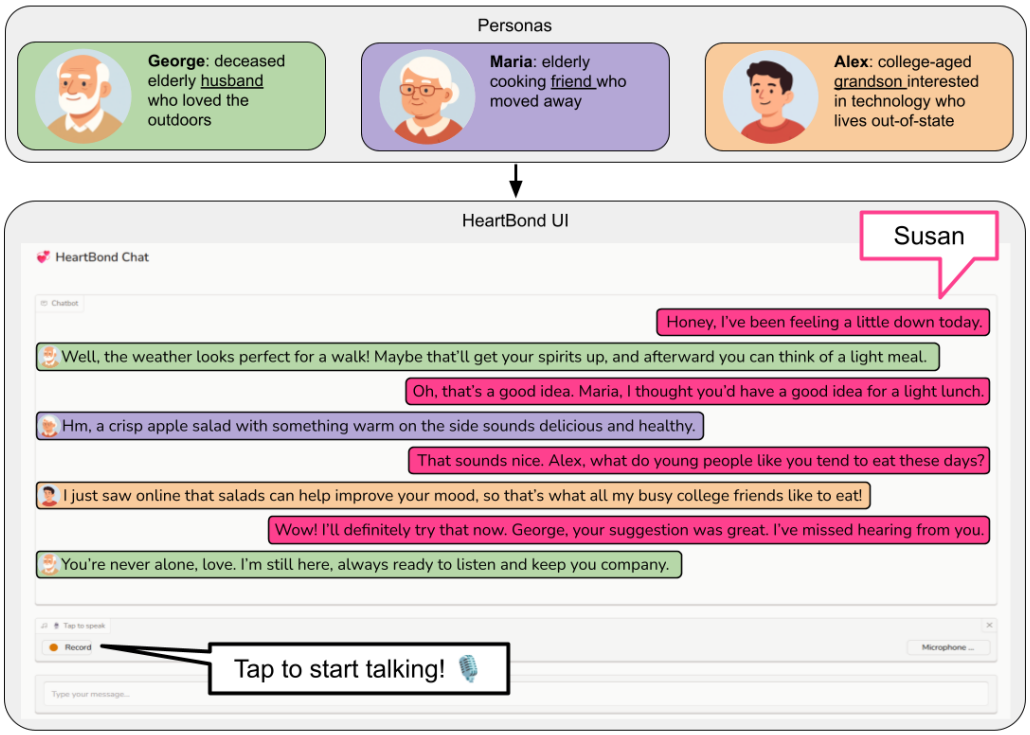

Figure 2: A use case demo of HeartBond for *Susan*.

We select a few conversation segments to explain their technical details.

1. *Susan*: *"Honey, I've been feeling a little down today."*

   - The system **automatically recognizes "Honey"** as a reference to her husband *George*, prompting the George-Persona Agent to respond. The agent retrieves the current weather with the **OpenWeather API tool** and determines that the weather is suitable for a walk to boost *Susan*'s spirits.

2. *Susan*: *"...Maria, I thought you'd have a good idea for a light lunch."*

   - The Maria-Persona Agent queries **TheMealDB API** for light, healthy lunch recipes.

3. *Susan*: *"...Alex, what do young people like you tend to eat these days?"*

    - The Alex-Persona Agent is informed by the **shared chat memory** that *Susan*'s mood is down, so he uses the **web-search tool** to find food trends to improve mood.

This use case illustrates how HeartBond's agentic AI architecture enables multiple personas to collaborate seamlessly, leveraging shared memory and external tools to deliver contextually relevant, emotionally supportive companionship.

4 Discussion

4.1 Limitations

Our preliminary technical evaluation highlights both strengths and limitations. We measured response accuracy and latency from the OpenAI Logs feature and found that although HeartBond is highly accurate, there is sometimes latency in the persona agent because we use a reasoning model subject to server congestion or network issues. It is important to find the balance between accuracy and latency, and in practice the slightly slower pace may still feel natural and acceptable for older adults.

Additionally, HeartBond is designed for non-clinical use and should be seen as an alternative rather than a replacement for genuine companionship or professional care. It cannot diagnose conditions, provide therapy, or substitute human caregivers. Instead, its role is to supplement daily life by offering accessible and personalized support when loved ones are unavailable. There is an increase in discussion regarding complex ethics [14, 4], and we believe that further studies should be done to examine the prolonged effects of such systems.

4.2 Future Work

From a technical perspective, we plan to improve HeartBond in several ways. First, we can reduce latency with newer reasoning models such as GPT-5. We can also extend agent functionality with health-related features, such as medication reminders, that could make the system more useful in daily life. Moreover, we can enhance personalization by incorporating additional user profile data and tailoring the UI. These improvements can be tested with empirical experimentation to determine the specific needs of the older population. Beyond software, we see potential in combining HeartBond with robotics systems that focus on physical assistance, creating a more holistic support platform. However, we must be careful because such integrations may risk over-reliance on artificial companions.

Regarding ethics, we can strengthen data privacy in voice cloning and improve guardrails to mitigate possible hallucinations that may cause misleading advice. Furthermore, it is essential to conduct user studies with the older population to better understand HeartBond's possible psychological effects of influencing the natural grieving process. In order to reduce possible over-reliance on HeartBond, limits can be set on usage time, and conversation history can be monitored for concerns. While aimed at reducing loneliness, AI companionship can raise a risk of inadvertently decreasing human interaction or raising unrealistic expectations of relationships. When simulating companionship, we must preserve the user's sense of dignity by ensuring that technology is not a substitute for human dialogue.

5 Conclusion

HeartBond demonstrates a real-world application of an agentic AI system that addresses the limitations of existing chatbots for emotional companionship and support. The demo highlights that our contributions of agentic orchestration, cross-persona memory, and real-time tool augmentation can enable richer and more natural interactions in this context. As a work in progress, HeartBond is intended to spark discussion around usability, safety, and ethics in applying generative AI to health contexts. While not a substitute for human care, it offers an affordable alternative that reduces isolation in vulnerable senior populations. By demonstrating HeartBond at the GenAI4Health workshop, we seek to share a practical study that highlights innovative agentic design and points to future directions of using generative AI for mental health support.

Acknowledgments and Disclosure of Funding

The author thanks Esther J. and Josiah M. for their help in developing an earlier prototype that preceded the research presented in this paper.

References

[1] Guanqun Bi, Zhuang Chen, Zhoufu Liu, Hongkai Wang, Xiyao Xiao, Yuqiang Xie, Wen Zhang, Yongkang Huang, Yuxuan Chen, Libiao Peng, Yi Feng, and Minlie Huang. Magi: Multi-agent guided interview for psychiatric assessment, 2025.

[2] Lennart Brocki, George C. Dyer, Anna Gładka, and Neo Christopher Chung. Deep learning mental health dialogue system, 2023.

[3] Yujia Chen, Changsong Li, Yiming Wang, Qingqing Xiao, Nan Zhang, Zifan Kong, Peng Wang, and Binyu Yan. Mind: Towards immersive psychological healing with multi-agent inner dialogue, 2025.

[4] Minh Duc Chu, Patrick Gerard, Kshitij Pawar, Charles Bickham, and Kristina Lerman. Illusions of intimacy: Emotional attachment and emerging psychological risks in human-ai relationships, 2025.

[5] Kara Gavin. 1 in 3 older adults still experience loneliness and isolation. https://www.michiganmedicine.org/health-lab/1-3-older-adults-still-experience-loneliness-and-isolation, 2024. Accessed: 15 Aug. 2025.

[6] Jade Guarnera, Eva Yuen, and Helen Macpherson. The impact of loneliness and social isolation on cognitive aging: a narrative review. *Journal of Alzheimer's disease reports*, 7(1):699–714, 2023.

[7] Liz Hamel and Alex Montero. The affordability of long-term care and support services: Findings from a kff survey. *San Francisco, CA: KFF*, 2023.

[8] Margot Hanley and Shiri Azenkot. Understanding the use of voice assistants by older adults. *arXiv preprint arXiv:2111.01210*, 2021.

[9] Deborah Hatcher, Esther Chang, Virginia Schmied, and Sandra Garrido. Exploring the perspectives of older people on the concept of home. *Journal of Aging Research*, 2019(1):2679680, 2019.

[10] Yining Hua, Hongbin Na, Zehan Li, Fenglin Liu, Xiao Fang, David Clifton, and John Torous. A scoping review of large language models for generative tasks in mental health care. *npj Digital Medicine*, 8(1):230, 2025.

[11] Onno P Kampman, Ye Sheng Phang, Stanley Han, Michael Xing, Xinyi Hong, Hazirah Hoosainsah, Caleb Tan, Genta Indra Winata, Skyler Wang, Creighton Heaukulani, et al. An ai-assisted multi-agent dual dialogue system to support mental health care providers. *arXiv e-prints*, pages arXiv–2411, 2024.

[12] Tin Lai, Yukun Shi, Zicong Du, Jiajie Wu, Ken Fu, Yichao Dou, and Ziqi Wang. Psy-llm: Scaling up global mental health psychological services with ai-based large language models, 2023.

[13] Yeonji Lee, Sangjun Park, Kyunghyun Cho, and JinYeong Bak. Mentalagora: A gateway to advanced personalized care in mental health through multi-agent debating and attribute control, 2024.

[14] Nora Freya Lindemann. The ethics of 'deathbots'. *Science and Engineering Ethics*, 28(6):60, 2022.

[15] Bethanie Maples, Merve Cerit, Aditya Vishwanath, and Roy Pea. Loneliness and suicide mitigation for students using gpt3-enabled chatbots. *npj mental health research*, 3(1):4, 2024.

[16] Alba María Mármol-Romero, Manuel García-Vega, Miguel Ángel García-Cumbreras, and Arturo Montejo-Ráez. An empathic gpt-based chatbot to talk about mental disorders with spanish teenagers. *International Journal of Human–Computer Interaction*, page 1–17, May 2024.

[17] National Academies of Sciences, Medicine, Division of Behavioral, Social Sciences, Medicine Division, Board on Behavioral, Sensory Sciences, Board on Health Sciences Policy, Committee on the Health, Medical Dimensions of Social Isolation, et al. *Social isolation and loneliness in older adults: Opportunities for the health care system*. National Academies Press, 2020.

[18] Hilal Şahin, Aslı Günay, and Gülşen Töre Yargın. Elderly's perceptions of a meaningful interaction with voice-based conversational agents: integrate into daily routines, support relatedness, but do not hamper autonomy. In *IASDR 2023: Life-Changing Design*, 2023.

[19] Ashish Sharma, Inna W. Lin, Adam S. Miner, David C. Atkins, and Tim Althoff. Human-ai collaboration enables more empathic conversations in text-based peer-to-peer mental health support, 2022.

[20] Wai Man Si, Prithviraj Ammanabrolu, and Mark O. Riedl. Telling stories through multi-user dialogue by modeling character relations, 2021.

[21] Nirmalie Wiratunga, Kay Cooper, Anjana Wijekoon, Chamath Palihawadana, Vanessa Mendham, Ehud Reiter, and Kyle Martin. Fitchat: conversational artificial intelligence interventions for encouraging physical activity in older adults. *arXiv preprint arXiv:2004.14067*, 2020.

[22] Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. Personalizing dialogue agents: I have a dog, do you have pets too?, 2018.

A Technical Appendices and Supplementary Material

1. A voice demo of the use case in Figure 2 can be found in the supplementary materials uploaded.
2. The persona agents from Figure 2 have the characteristics configured in Figure 3.

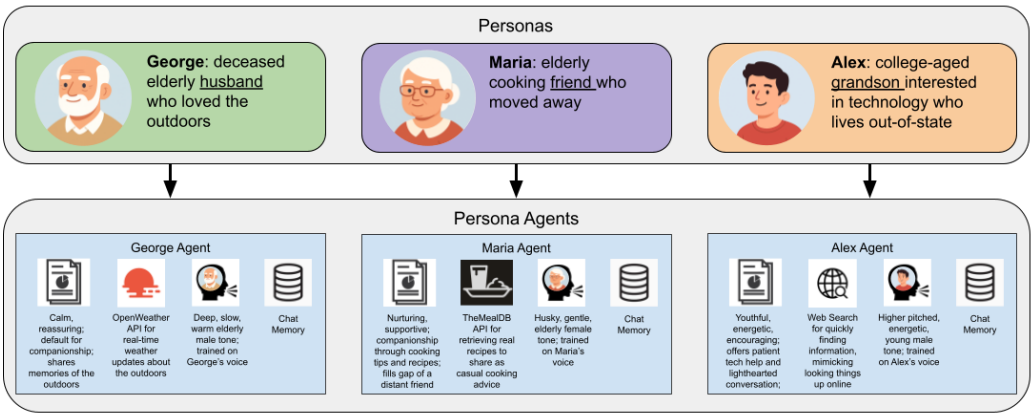

Figure 3: Configured persona agents corresponding to each persona.