Chat4Elderly: A Multi-Agent System for Personalized Wellness Using Generative AI and Wearable Technology

Demonstration Track

Vítor Crista
GECAD, Polytechnic of Porto
Porto, Portugal
vvrpc@isep.ipp.pt

Diogo Martinho
GECAD, Polytechnic of Porto
Porto, Portugal
diepm@isep.ipp.pt

Goreti Marreiros
GECAD, Polytechnic of Porto
Porto, Portugal
mgt@isep.ipp.pt

ABSTRACT

This demo presents Chat4Elderly, a Multi-Agent System (MAS) designed to support the well-being of elderly users through personalized and proactive interactions. The system combines Large Language Models (LLMs) with smartwatch sensor data to provide real-time, context-aware responses that address both mental and physical health needs. By analyzing conversational patterns and physical activity levels, it adapts to user preferences, offering personalized assistance and engagement. Over time, the system improves its interactions using stored knowledge, increasing personalization and supporting long-term well-being. This approach helps reduce loneliness and enhance quality of life (QoL), creating a more supportive and engaging experience for elderly users.

KEYWORDS

Conversation System, Generative AI, Multi-Agent Systems, Wearable Technology

ACM Reference Format:
Vítor Crista, Diogo Martinho, and Goreti Marreiros. 2025. Chat4Elderly: A Multi-Agent System for Personalized Wellness Using Generative AI and Wearable Technology: Demonstration Track. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 3 pages.

1 INTRODUCTION

As the global population ages, improvements in healthcare and living standards have led to a growing number of elderly people, creating new challenges that require innovative solutions [11]. Older adults face mobility problems that increase the risk of accidents, while psychological struggles such as loneliness, depression, and anxiety are common, all contributing to a decline in QoL [6, 13]. Technology presents a promising solution to these challenges as recent advances in generative AI [5], wearable technology [9] and MAS [9] have allowed the development of continuous monitoring systems that significantly improve elderly care. LLMs such as GPT [1] and Llama [14] models perform exceptionally well on a variety of general-purpose tasks, including answering questions, summarizing texts and generating code, consistently delivering impressive performance [15]. These technologies are now more accessible, driving innovation, and enabling customized solutions across multiple industries. However, their effectiveness can be limited in scenarios that require highly specific responses or adapted personalization for individual users. This is where techniques such as fine-tuning [8] and prompt engineering [12] become essential. Using customized prompts can optimize the performance of LLMs for specific use cases, such as improving elderly care [3]. This demo presents a MAS that combines generative AI and wearable technology to provide personalized and proactive interactions. The demonstration video can be found at: https://www.gecad.isep.ipp.pt/chat4elderly-backend

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

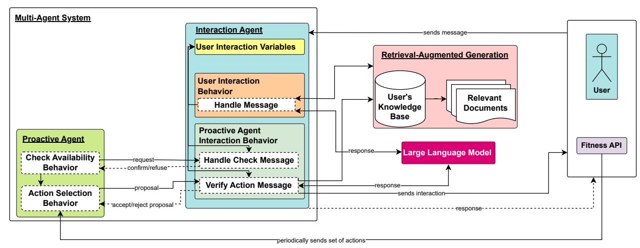

Figure 1: System Architecture

2 THE PROPOSED SYSTEM

The proposed system combines generative AI with wearable technology to create personalized and proactive interventions aimed at improving the well-being of elderly individuals. It uses LLMs to facilitate communication through two techniques. The first, retrieval-augmented generation (RAG) [7], retrieves relevant information from a database of past interactions, improving response accuracy and relevance. The second technique uses Fitbit smartwatch sensor data to assess the user's physical state in real time. By incorporating both sources of information, the system generates more accurate and contextually appropriate responses. Figure 1 shows the system architecture and explains how it functions. The system consists of three main components: MAS, RAG, and LLM. MAS includes two agents: the interaction agent, which manages communication, and the proactive agent, which processes sensor data and alerts the interaction agent when new actions are needed. The RAG stores and retrieves previous knowledge from past conversations, while the LLM generates the final message sent to the user.

2.1 Interaction Agent

The interaction agent serves as an intermediary, facilitating communication between the user and the proactive agent. To achieve this, the agent presents two main behaviors: the user interaction behavior and the proactive agent interaction behavior. Using the user interaction behavior, the agent processes incoming user messages and attempts to retrieve relevant past conversations from the knowledge base. To assess the relevance of previous conversations, a cosine similarity search is utilized, which compares the embeddings of previously conversations with those of the current user message, generated using the *all-minilm-l6-v2* embedding model. If RAG identifies relevant documents, they are appended to the conversation's system prompt, and the final model, *meta-llama-3.1-8b-instruct*, is invoked with the improved prompt. This model was chosen because it demonstrated the best balance between LLM quality and computational resource available [4, 10].The proactive agent interaction behavior involves engaging with the agent using messages that evaluate user interaction and trigger system actions.

2.2 Proactive Agent

The proactive agent communicates directly with the interaction agent. It fetches walking steps data from the Fitbit smartwatch sensor via API at the beginning of each day and processes it using the following algorithm.

Algorithm 1 Calculate User Profile

```
function ANALYZE_STEPS(steps)
    step_threshold ← 5000
    total_steps ← 0
    initialize avg_steps, most_active, least_active as empty lists
    for each day from 0 → 6 do
        day_steps ← sum(steps[day])
        total_steps ← total_steps + day_steps
        add day_steps to avg_steps
        periods ← { sum(steps[day][6:12]), sum(steps[day][12:18]),
sum(steps[day][18:24]), sum(steps[day][0:6]) }
        labels ← { 0, 1, 2, 3 }
        add labels[index_max(periods)] to most_active
        add labels[index_min(periods)] to least_active
    end for
    avg_daily_steps ← total_steps / 7
    for each daily_steps in avg_steps do
        if daily_steps ≥ step_threshold then
            step_status ← 1
        else
            step_status ← 0
        end if
    end for
    return avg_daily_steps, most_active, least_active
end function
```

The algorithm analyses data from the past seven days to identify activity patterns. Each day, it calculates the number of steps taken and divides them into four periods: morning, afternoon, evening, and night. It determines the periods with the highest and lowest activity levels and computes the average daily step count for the week. A threshold of 5000 steps [2] is used to evaluate the physical activity level, and the algorithm returns this evaluation along with the times of highest and lowest activity. The system only determines the specific hour and type of interaction to initiate a conversation, without specifying how to begin or what content to include. This is where generative AI, powered by LLMs, plays a crucial role. Figure

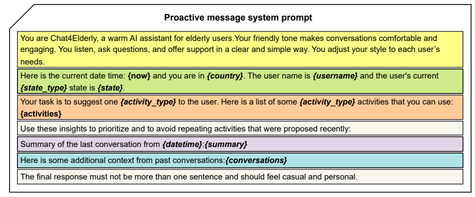

Figure 2: Prompt for proactive interaction

2 illustrates an example of the prompt used to generate a proactive interaction. This prompt is designed to start proactive conversations and it is divided into seven sections. The first section outlines the model's role. The second section is dynamic, including user details, location, and the current date and time for time-sensitive responses. The third section instructs the model to choose an appropriate interaction from a list based on context. The fourth section directs the model to avoid repeating previously suggested interactions, using past conversation data. The fifth section offers a summary of the last conversation for additional context. The sixth section includes relevant past interactions to inform the current task. The prompt concludes with a request for the model to provide one clear, concise suggestion to the user.

2.2.1 User Profiles. During registration, users provide details such as name, age, activity level, emotional state, and preferences. This data creates an initial profile, allowing system access without a smartwatch. Once connected, the system compares self-reported activity with Fitbit data for more accurate fitness assessments. Fitness data is not stored but retrieved daily to adjust user levels. Only profiles and conversations are stored, and users can delete their account and reset personalization. The system adjusts prompts based on emotional state and uses sentiment analysis to enhance proactive interactions.

3 CONCLUSION

This paper addresses the challenges of aging, such as physical mobility and social isolation, which negatively impact the QoL. It proposes Chat4Elderly, a MAS system that integrates AI technologies, particularly generative AI, to provide dynamic, personalized support for elderly individuals, improving both physical and mental health through conversation-based systems. While the current system is ready for testing with real users, future work will focus on improving metrics and personalization by collecting more data on elderly users' daily activities. This approach aims to promote active and healthy aging, supporting a balanced physical, mental, and social well-being.

ACKNOWLEDGMENTS

This research work was developed under the project RM4Health (ITEA-2021-21022-RM4Health), funded by the European Regional Development Fund (ERDF) within the project number COMPETE2030-FEDER-00391100, and funded by National Funds through the Portuguese FCT — Fundação para a Ciência e a Tecnologia under the R&D Units Project Scope, UIDB/00760/2020 (https://doi.org/10.54499/UIDB/00760/2020).

REFERENCES

[1] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774* (2023).

[2] T Aranyavalai, C Jalayondeja, W Jalayondeja, S Pichaiyongwongdee, J Kaewkungwal, and JJ Laskin. 2020. Association between walking 5000 step/day and fall incidence over six months in urban community-dwelling older people. *BMC geriatrics* 20 (2020), 1–11.

[3] Jin Chen, Zheng Liu, Xu Huang, Chenwang Wu, Qi Liu, Gangwei Jiang, Yuanhao Pu, Yuxuan Lei, Xiaolong Chen, Xingmei Wang, et al. 2024. When large language models meet personalization: Perspectives of challenges and opportunities. *World Wide Web* 27, 4 (2024), 42.

[4] Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783* (2024).

[5] Kathleen Fear, Conrad Gleber, et al. 2023. Shaping the future of older adult care: ChatGPT, advanced AI, and the transformation of clinical practice. *JMIR aging* 6, 1 (2023), e51776.

[6] Eva Gerino, Luca Rollè, Cristina Sechi, and Piera Brustia. 2017. Loneliness, resilience, mental health, and quality of life in old age: A structural equation model. *Frontiers in psychology* 8 (2017), 2003.

[7] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems* 33 (2020), 9459–9474.

[8] Wei Lu, Rachel K Luu, and Markus J Buehler. 2024. Fine-tuning large language models for domain adaptation: Exploration of training strategies, scaling, model merging and synergistic capabilities. *arXiv preprint arXiv:2409.03444* (2024).

[9] Solange Mendes, Jonas Queiroz, and Paulo Leitão. 2017. Data driven multi-agent m-health system to characterize the daily activities of elderly people. In *2017 12th Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, 1–6.

[10] John Mendonça, Alon Lavie, and Isabel Trancoso. 2024. On the Benchmarking of LLMs for Open-Domain Dialogue Evaluation. *arXiv preprint arXiv:2407.03841* (2024).

[11] Hannah Ritchie, Lucas Rodés-Guirao, Edouard Mathieu, Marcel Gerber, Esteban Ortiz-Ospina, Joe Hasell, and Max Roser. 2024. Population growth. *Our World in Data* (2024).

[12] Sander Schulhoff, Michael Ilie, Nishant Balepur, Konstantine Kahadze, Amanda Liu, Chenglei Si, Yinheng Li, Aayush Gupta, HyoJung Han, Sevien Schulhoff, et al. 2024. The Prompt Report: A Systematic Survey of Prompting Techniques. *arXiv preprint arXiv:2406.06608* (2024).

[13] Guddo Sharma and Sandeep Kumar Morishetty. 2023. Common mental and physical health issues with elderly: A narrative review. *ASEAN Journal of Psychiatry* 23 (2023), 1–11.

[14] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971* (2023).

[15] Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, et al. 2023. A survey of large language models. *arXiv preprint arXiv:2303.18223* (2023).