Symbiosis and Synesthesia in Proactive Conversational Agents for Healthy Ageing

Mario Alessandro Bochicchio[1*], Simona Corciulo[2]

[1]*Department of Computer Science, University of Bari Aldo Moro, Italy*
[2]*Department of Computer Science, University of Turin, Italy*

Abstract
The aging of the population is an important and unprecedented phenomenon of the 21st century that pushes for innovative solutions to delay the psychophysical decline of the elderly, maintain as much as possible an acceptable level of autonomy and a good degree of socialization, prevent where possible acute events and reduce hospitalization. In this context, the paper addresses the development of a conversational companion for elderly individuals as part of the Age-It Project. The aim is to support active and healthy aging through symbiotic AI concepts. The project involves small-scale testing of proactive conversational agents to monitor health, provide cognitive stimulation, reduce loneliness, and inform caregivers about patients' health status. The requirements and system architecture are discussed, emphasizing the importance of ethical considerations and personalized interactions. The paper also explores the potential of combining AI with the concept of synesthesia to enhance empathy and effectiveness in care. The prototype system has been implemented and tested in the laboratory, with plans for further development and benchmarking.

1. Introduction

Population ageing is a major, and unprecedented, 21st century phenomenon which, as a trend, concerns the whole world. Adopting the World Health Organization guidelines, the Age-It project (https://ageit.eu/wp/), co-funded by the European Union - Next Generation EU Programme, addresses this situation by promoting the adoption of policies and strategies based on the “active and healthy ageing” framework [1]. As part of the Age-It project, the Spoke 8 of Age-It aims to support the needs and strengthen abilities to foster the health of older patients, design and adopt innovative technologies in different life domains and environments and create a large dataset that can be exploited also through AI models to develop personalized intervention.

Spoke 8 also involves small-scale testing of proactive conversational agents that should support patients with cognitive stimulation, reduce feelings of loneliness, help them comply with medical prescriptions, and inform relatives and caregivers about their health status and possible emergencies.

This paper, starting with an analysis of related work (Section 2), discusses the requirements that such a conversational agent should meet (Sec. 3), the reasons why it should be deeply rooted on symbiotic AI concepts and the close links with the concept of synesthesia (Sec. 4), a preliminary definition of the system architecture, and some thoughts on implementation aspects (Sec. 5).

2. Related Works

Conversational Agents (CAs in the following) dedicated to personalized interventions, particularly in the context of wellness, frequently face the challenge of having to operate within predefined and limited dialogue patterns and having to handle circumscribed and anticipated user input [2]. Despite the potential, the use of CAs is still in its early stages, underscoring the urgent need for further research to ensure development that is both safe and effective. It is critical to recognize the inherent limitations in CAs, including the propensity for bias and the dissemination of untrue information, and to consider related ethical issues, particularly privacy concerns. These issues become evident, for example, in the design of conversational companions created to mitigate problems such as loneliness and social isolation, even in response to conditions such as depression, where deeply personal interaction requires careful consideration of ethical implications [3] [4].

This focus on the efficacy and safety aspects of interaction marks a shift toward more sophisticated and potentially more invasive forms of care. Li et al. [5] conducted a prospective trial involving 278 patients with osteoarthritis and sarcopenia observing how the synergistic use of ChatGPT-4 and wearable devices not only facilitated access to care but also significantly improved the quality of rehabilitation care, showing a progression toward more personalized and impactful interventions.

This technological advancement, however, brings with it new challenges, as demonstrated by the comparative evaluation of LLM-based chatbots in Alzheimer's recognition [6]. While Bard excelled in disease identification, it showed a tendency to overestimate the presence of disease, unlike GPT-4, which stands out for its accuracy in recognizing cognitively healthy subjects. Overall, to date, it is difficult for a chatbot to reach the levels required for clinical applications, although some preliminary evaluations point to the great potential of CAs in the diagnostic field [7].

3. A Proactive Conversational Companion for Elderly People

Spoke 8 of the Age-It project involves a multicomponent intervention on frail people older than 65 years. The multicomponent intervention includes moderate-intensity physical activity, personalized nutritional counseling, active aging health education, and social recreational activities (e.g., singing, dancing, music etc.). Downstream of numerous meetings conducted with multidisciplinary experts (physicians, pedagogists, psychologists, computer scientists, data scientists, sociologists, ...) from January 2023 to January 2024, among other activities, the project defined small-scale testing of proactive conversational agents capable of:

1) Collect patient health status monitoring data (temperature, sleep/wake time, heart rate, blood oxygenation level, etc.) and data related to the performance of Activities of Daily Life (ADLs) (e.g., posture, speed of movement, amount of physical activity, etc.)
2) Vocally interacting with project participants by providing stimuli on activities to be performed: e.g., adherence to treatment (i.e., taking prescribed medication at the

right time), physical exercises, cognitive stimuli, invitation to communicate with friends or family members, etc.

3) Interact vocally with the patient's family members or caregivers, informing them in case of an emergency (e.g., patient's illness, abnormal or risky behavior, etc.);
4) Interact vocally with medical personnel, summarizing the salient elements related to overall health status (improving, stable, worsening), and reporting specific events (abnormal temperature, poor sleep quality, failure to take medication, falls, etc.) useful for diagnostic evaluation and possible therapeutic upgrade or revision.

The requirements thus described were collected in the form of Use Cases and Sequence Diagrams, represented in UML [8], discussed and fine-tuned with the specialists participating in the project. In Q3 and Q4 2024, in order to take into account the opinion of direct stakeholders, it is planned to review and extend the requirements together with representatives of associations of patients and family members of frail individuals through focus groups.

Based on the requirements thus defined and the available scientific literature [9] [10] it became apparent that conventional Natural Language Processing (NLP) approaches and current LMMs, taken individually, do not meet the needs summarized in points 1 to 4.

For this reason, we defined the architecture described in Fig. 1 that integrates the conversational and reasoning capabilities of an LMM with a Retrieval Augmented Generation (RAG) system capable of storing and, subsequently, extracting from a vector archive and synthesizing personal information related to a patient's health status and its temporal evolution.

4. Symbiosys, Synesthesia, and Companionship for the Elderly

The dialogue between the physician and the patient has been defined as “the most powerful, sensitive, and most versatile instrument available to the physician” [11]. It is fundamental to effective and empathic care; it goes beyond history and diagnosis in that it establishes the rapport and trust necessary to address health needs. Physicians possess considerable skills in history taking and the broader "diagnostic dialogue," but access to these skills remains episodic [12] for reasons that can be mitigated by an appropriate usage of AI systems, as discussed in a recent work sponsored by Google Research and Google DeepMind [13]. In the paper, the authors discuss the results of a structured objective clinical examination with validated simulated patients interacting with a conversational AI or primary care physicians via a text interface, finding that AI is superior to primary care physicians in empathy, treatment plan management and other aspects.

Given the importance that clinicians attribute to aspects such as empathy and trust, it is reasonable to assume that different patients with different diseases and personal histories may produce different mental states and may have different expectations in relation to conversational AI such as the one described. Thus, especially in the hypothesis of Conversational Companions to be placed side by side for long periods of time (months or years) with frail or chronically ill patients, it is reasonable to assume that the adoption of symbiotic AI principles may further enhance the long-term AI-patient relationship by allowing the AI to combine

generative capabilities with the ability to make assumptions about the patient's mental state and expectations.

To better understand the profound impact that the concept of symbiosis can have in Conversational Companions in clinical settings, it seems appropriate to retrace the basic steps of its evolution. The concept of symbiosis, proposed by the German botanist and mycologist Anton de Bary in 1878, defines exchange relationships between living organisms, classifying them as forms of mutualism (where both partners benefit from the relationship) or parasitism (where one exploits the other for its own benefit). In his pioneering publication "Cybernetics: Or Control and Communication in the Animal and the Machine" in 1948, Norbert Wiener uses the concept of symbiosis to characterize the integrative relationship between human and machine, emphasizing the profound differences from the more basic and imprecise concepts of use (of machines) or synergy (with machines) and foreshadowing, for machines, the development of "predictive" capabilities with respect to the needs and expectations of the human symbiont. In human subjects, this predictive ability is associated with the concept of "theory of mind" and refers to the ability to attribute mental states to oneself and others, and to understand that others have mental states different from ours.

The theory of mind for Baron-Cohen [14] is a "lens" through which we view the social world, and which allows us to interpret and predict the behavior of others, while for Dennett [15] this allows us to see others not only as physical entities, but as complex beings with desires, beliefs, and intentions.

Returning, then, to Conversational Companions for clinical use, it now appears evident that since they must be equipped with empathic sensitivities, credibility, and the ability to adapt their responses to the mental state of fragile subjects, they must be able to assess the needs, expectations, and mental states of the human symbiont and, specifically, of fragile subjects with potential cognitive problems and problematic emotional states.

To develop a Conversational Companion for frail individuals that could prospectively include these features, the authors also made use of literature findings related to a specific rare anomaly of the human sensory-cognitive system, termed "mirror-touch synesthesia" and closely related to empathy [16] and hypochondria. In mirror-touch synesthesia, which affects about 1.6% of the population, synesthetes experience tactile sensations on their bodies when they see other people being touched [17]. Incorporating empathy models based on this concept into Conversational Companions for Elderly could significantly enrich the ability of machines to "describe and process" human emotions. Such a conversational clinical AI, by accompanying a specific frail person in his or her daily activities, could over time tune its theory of mind to the characteristics and expectations of the specific frail person with whom it accompanies, potentially enabling more effective and personalized interactions, but detailed discussion of this is beyond the scope of this paper.

5. Preliminary Design of a Proactive Conversational Agent for Multicomponent Intervention in the age-It Project

In Fig. 1 is depicted the logic schema of a very first iteration of the above-described Conversational Companion for elderly people. For sake of simplicity, on the right side of Fig.1 only the patients are indicated but, as anticipated in Section 3, the Companion is also able of alerting the patient's relatives/caregivers in case of problems, answering their question about the patient's health status, and informing physicians about the patient's health/wellness in a the time interval of interest.

The main logic blocks represented in the schema are:

- a proactive agent, based on a program-driven and event-driven logic, able to activate the Conversational Companion when needed, independently from the question- answer mechanism usually adopted by LLMs. The Proactive Agent is also able to activate/deactivate specific health-monitoring devices, not shown in Fig.1 for sake of clarity (e.g. wearable wristbands, medical sensors, environmental sensors, etc.) and store monitoring data in the vector store or structured-data store (not shown in figure).
- a Retrieval Augmented Generation (RAG) subsystem, based on the Large Language Model (ChatGPT 4 at the moment), a purposely-defined logic to generate the embedding associate to each content stored in the Vector Store (the OpenAI embedding model, not represented in Fig.1, is used at the moment), and the Vector Store (the Chroma Vector Store, in the current version). The RAG is in charge to search, extract and transform into meaningful text messages all contents stored in the vector store or on the Web, when needed.
- The "Experts" collection, i.e. the set of documents, guidelines, best practices and specific medical prescriptions, stored in the Vector DB, describing the multicomponent interventions associated with the specific pathologies of the specific patient. This collection is used by the proactive agent to decide when a specific therapeutic action or cognitive stimulus or other intervention must be activated.
- The "Personal Data" collection, containing the relevant events, health-monitoring data, actions, environmental data, etc. collected by the Proactive Agent and stored in the Vector DB.
- a text to speech (LabEleven in this version) and a speech to text (OpenAI Whisper) to support a natural vocal interaction among the system and the different users (the patient, the caregivers, the relatives, the physicians etc.).

As of the first quarter of 2024, the system has been implemented in a preliminary version and tested in the laboratory for major functionalities. Due to the limitations of the Chroma Vector Store, the "expert" collection data is only a subset of the data needed for the experimental phase. In addition, the current implementation is not suitable for experiments with real patients because of the lack of privacy and anonymity inherent in the adoption of ChatGPT 4. However, the system can be used for simulations based on realistic anonymized data downloaded from Physionet and Kaggle, pending a fully private and local version based on LLama 70b and not connected to the Web.

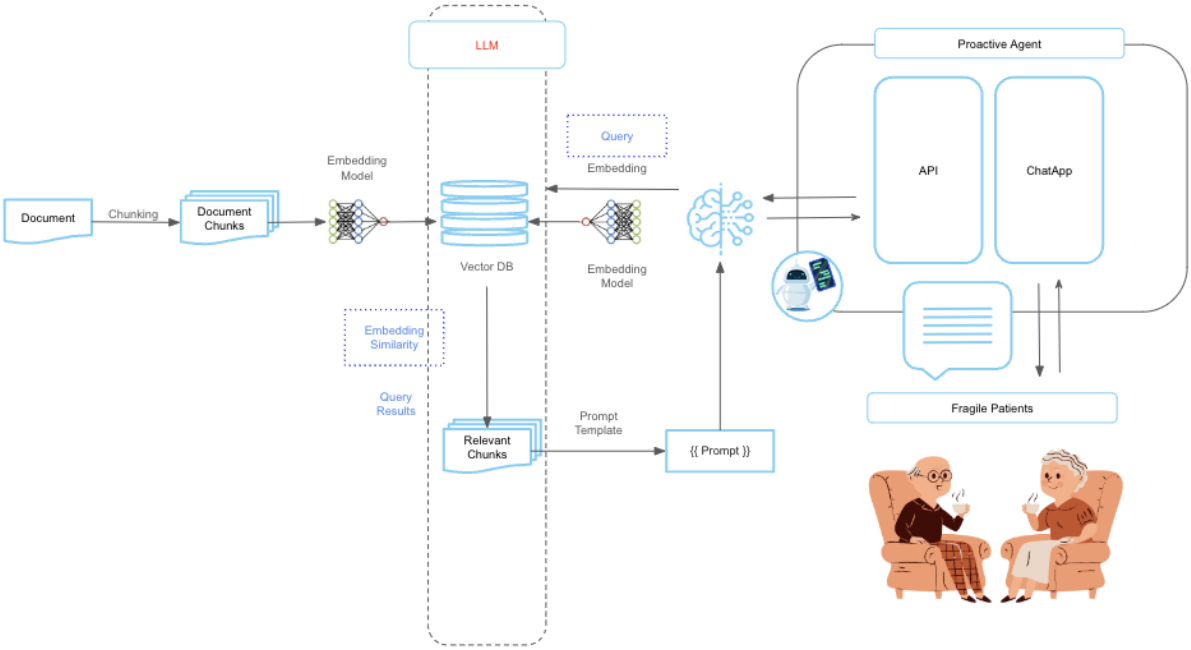

Figure 1: Logic Schema of the Conversational Companion for Elderly.

6. Conclusions

The paper defines the main requirements for a Conversational Companion for elderly and frail patients, to be used in the Age-It PNRR Project - Spoke 8 - Multicomponent Intervention.

The system prototype has been designed, implemented, tested, and is currently used with anonymous data for simulation, demonstration, extension and fine tuning purposes.

In the next steps we plan to design and build a more complete implementation of the system and benchmark his “theory of mind” features, toward more empathic and effective Conversational companions for clinical applications. To test the system on the field, with real patients, an instance of Llama 70b will be run on a Mac M3 Pro. The system will be connected to a Local Area Network including all needed sensors and devices, but not connected to Internet to preserve the privacy and confidentiality of data collected by the patient. The experiment will undergo the local Ethical Committee for authorization and will implement the prescribed privacy, security and confidentiality measures.

Acknowledgments

This publication was partially funded by the European Union - Next Generation EU, in the context of the National Recovery and Resilience Plan, Investment Partenariato Esteso PE8 "Conseguenze e sfide dell'invecchiamento", Project Age-It (Ageing Well in an Ageing Society), CUP: B83C22004800006.

References

[1] https://iris.who.int/handle/10665/186463

[2] Martins, A., Nunes, I., Lapão, L., & Londral, A. (2024). Unlocking Human-Like Conversations: Scoping Review of Automation Techniques for Personalized Healthcare Interventions using Conversational Agents. International Journal of Medical Informatics, 105385

[3] Alessa, A., & Al-Khalifa, H. (2023, July). Towards designing a ChatGPT conversational companion for elderly people. In Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments (pp. 667-674)

[4] Javaid, M., Haleem, A., & Singh, R. P. (2023). ChatGPT for healthcare services: An emerging stage for an innovative perspective. BenchCouncil Transactions on Benchmarks, Standards and Evaluations, 3(1), 100105

[5] Li et al. (2023). ChatGPT-4 and Wearable Device Assisted Intelligent Exercise Therapy for Co-existing Sarcopenia and Osteoarthritis (GAISO): A feasibility study and design for a randomized controlled PROBE non-inferiority trial

[6] Chen, J. M. (2024). Performance Assessment of ChatGPT vs Bard in Detecting Alzheimer's Dementia. arXiv preprint arXiv:2402.01751

[7] Tu, T., Palepu, A., Schaekermann, M., Saab, K., Freyberg, J., Tanno, R., Wang, A., Li, B., Amin, M., Tomašev, N., Azizi, S., Singhal, K., Cheng, Y., Hou, L., Webson, A., Kulkarni, K., Mahdavi, S.S., Semturs, C., Gottweis, J., Barral, J., Chou, K., Corrado, G.S., Matias, Y., Karthikesalingam, A., & Natarajan, V. (2024). Towards Conversational Diagnostic AI. ArXiv, abs/2401.05654

[8] https://www.omg.org/spec/UML/2.5.1/PDF

[9] Alessa, A., & Al-Khalifa, H. (2023, July). Towards designing a ChatGPT conversational companion for elderly people. In Proceedings of the 16th International Conference on PErvasive Technologies Related to Assistive Environments (pp. 667-674)

[10] Chen, Z., Wu, J., Zhou, J., Wen, B., Bi, G., Jiang, G., ... & Huang, M. (2024). ToMBench: Benchmarking Theory of Mind in Large Language Models. arXiv preprint arXiv:2402.15052

[11] Engel, G. L. & Morgan, W. L. Interviewing the patient (1973)

[12] Rennie, T., Marriott, J. & Brock, T. P. Global supply of health professionals. N Engl J Med 370, 2246–7 (2014)

[13] Tu, T., Palepu, A., Schaekermann, M., Saab, K., Freyberg, J., Tanno, R., Wang, A., Li, B., Amin, M., Tomašev, N., Azizi, S., Singhal, K., Cheng, Y., Hou, L., Webson, A., Kulkarni, K., Mahdavi, S.S., Semturs, C., Gottweis, J., Barral, J., Chou, K., Corrado, G.S., Matias, Y., Karthikesalingam, A., & Natarajan, V. (2024). Towards Conversational Diagnostic AI. ArXiv, abs/2401.05654

[14] Baron-Cohen, S. (1997). Mindblindness: An Essay on Autism and Theory of Mind. Cambridge, MA: MIT Press

[15] Dennett, D. C. (1987). The Intentional Stance. Cambridge, MA: MIT Press

[16] Banissy, M., Ward, J. Mirror-touch synesthesia is linked with empathy. Nat Neurosci 10, 815–816 (2007). https://doi.org/10.1038/nn1926

[17] Sathian, K., & Ramachandran, V. S. (Eds.). (2019). Multisensory perception: From laboratory to clinic. Academic Press