Enhancing Cognitive Therapies for Older Adults with EBO: An Autonomous Social Robot System Integrating LLMs and Therapist-in-the-Loop

Antonio Blanco-Castillo, Jackie Lee Chong Ojeda, Alicia Condón, and Pedro Núñez[0000−0002−3615−8833]

Universidad de Extremadura, Spain pnuntru@unex.es
http://robolab.unex.es

Abstract. The growing population of older adults necessitates effective cognitive therapies. Social robots have potential as therapeutic tools, providing companionship, entertainment, and support. Traditional methods for cognitive engagement need more in terms of personalization and interaction. This paper presents the development of a system using the social robot EBO for interactive cognitive games and therapies, leveraging Large Language Models (LLMs). The system integrates the narrative generation via Llama on the CORTEX architecture, user response capture with Whisper, and narration through the MeloTTS engine. Tailored for older adults with mild to moderate cognitive impairment, it personalizes interactions based on individual interests, preferences, and cognitive levels. The therapist-in-the-loop approach allows the therapist to participate in the initial game configuration. The system conducts the therapy sessions autonomously, providing a final summary for analysis. This personalized therapy aims to enhance user experience and maximize the therapeutic utility of the system in promoting cognitive and emotional well-being within this demographic group. We evaluated various open-source LLMs and demonstrated their potential effectiveness through a pilot study, showcasing enhanced user experience. This approach could significantly enhance therapies for older adults using social robots as interactive tools.

Keywords: Social robots · Cognitive therapies · Older adults care · LLM.

1 Introduction

As the global population ages, the prevalence of cognitive impairments among older adults continues to rise, posing significant challenges for healthcare systems worldwide. Cognitive impairments, which range from mild cognitive decline to severe dementia, impact millions of elderly adults, reducing their quality of life and increasing the burden on caregivers and healthcare providers. Addressing these challenges requires innovative approaches that leverage technological advancements to provide effective, personalized, and scalable cognitive therapies.

Social robots have emerged as promising tools in cognitive therapy for older adults. These robots, designed to interact socially with humans, offer companionship, entertainment, and therapeutic support [10]. They can play a crucial role in cognitive engagement through various activities, including interactive storytelling, memory games, and other cognitive exercises. However, traditional methods of cognitive engagement often fall short in terms of personalization and interaction, which are critical for maintaining the interest and motivation of users with cognitive impairments.

This paper presents the development of an advanced system utilizing the social robot EBO for delivering interactive cognitive games and therapies. EBO, originally designed to promote computational thinking in educational settings, has been significantly upgraded to meet the current needs of cognitive therapy. This includes the integration of the CORTEX architecture [5], a robust cognitive framework that supports the deployment of various AI agents to enhance the robot's interactive capabilities. Our system employs Large Language Models (LLMs) such as Llama [12] for generating engaging narratives and interactions. Additionally, it incorporates Whisper [9] for capturing and transcribing user responses and the MeloTTS engine [14] for synthesizing natural-sounding speech. These technologies enable EBO to deliver personalized and adaptive cognitive therapies tailored to the individual interests, preferences, and cognitive levels of older adults with mild to moderate cognitive impairments.

A key feature of our approach is the "therapist-in-the-loop" methodology. This involves the therapist in the initial configuration of the cognitive games and therapies, ensuring that the activities are aligned with each user's specific needs and capabilities. Once configured, the system operates autonomously, conducting therapy sessions and providing a comprehensive summary of the interactions for further analysis by the therapist. This approach enhances the therapy's personalization and maximizes its therapeutic utility by allowing continuous adaptation and improvement based on user feedback and performance.

The primary aim of this system is to enhance user experience and promote cognitive and emotional well-being within the target demographic group. By leveraging advanced AI technologies and a user-centered design, we aim to provide a solution that is both effective and enjoyable for older adults, helping to mitigate the effects of cognitive decline and improve their overall quality of life.

In this paper, we will detail the development and implementation of the EBO-based system, discussing its architecture, functionalities, and the specific AI techniques employed. We will also present the results of evaluating different LLM models and a pilot study conducted with testers and caregivers, demonstrating the system's effectiveness and potential as a therapeutic tool.

2 Related works

The utilization of social robots in cognitive therapy for the elderly has gained significant attention in recent years due to the increasing prevalence of cognitive impairments among older adults. Social robots have been developed to provide mental support, companionship, and cognitive stimulation to elderly individuals, particularly those with dementia. For instance, robots like Paro and other interactive devices have demonstrated effectiveness in enhancing the emotional well-being of elderly patients through social interaction and cognitive engagement. These robots can support various therapeutic activities, such as memory games and storytelling, which are crucial for maintaining cognitive functions. However, challenges remain in customizing these interactions to cater to the individual needs of users, as traditional approaches often lack the necessary personalization and adaptability required for diverse cognitive impairments [11]. Additionally, studies have shown that social robots can reduce stress and improve social interaction among elderly individuals [1]. Novel robots like Rassle have been developed to engage elderly users through touch-based interactions, improving their quality of life [15]. Another study highlights the differential efficacy of social robots based on various neuropsychiatric profiles, demonstrating significant improvements in emotional well-being [3].

Automatic Speech Recognition (ASR) and Text-to-Speech (TTS) technologies significantly impact the effectiveness of interactions between social robots and elderly users. ASR enables robots to understand and respond to spoken language, making interactions more intuitive and user-friendly. On the other hand, TTS allows robots to communicate naturally with users by generating human-like speech. Studies have shown that integrating ASR and TTS technologies into social robots enhances their ability to engage elderly individuals by providing clear, understandable, and interactive communication. This integration helps reducing the cognitive load on users, facilitates smoother interactions, and makes the therapeutic experience more engaging and effective [7]. Moreover, using TTS with emotional recognition capabilities allows robots to adapt their responses based on the user's emotional state, enhancing the overall interaction quality [4]. While there are challenges in ensuring the naturalness and accuracy of ASR and TTS technologies, advancements in these areas continue to improve the effectiveness of social robots in therapeutic settings.

Comprehensive systems that combine ASR, TTS, and Large Language Models (LLMs) are emerging as powerful tools in cognitive therapy for the elderly. These integrated systems leverage advanced AI to provide personalized and adaptive interactions, offering a more tailored therapeutic experience. For example, combining ASR, TTS, and LLMs such as GPT-3 enables robots to generate meaningful and contextually relevant responses, enhancing engagement and cognitive stimulation. Such systems have been shown to maintain user interest over extended periods and provide more effective cognitive engagement than traditional methods. These advancements address limitations in previous robotic systems, which often struggled with maintaining long-term engagement due to repetitive interactions [13]. Additionally, personalized cognitive stimulation programs that leverage AI have been proven to maintain normal cognitive functioning and delay cognitive decline, demonstrating the potential of AI-enhanced therapy [6].

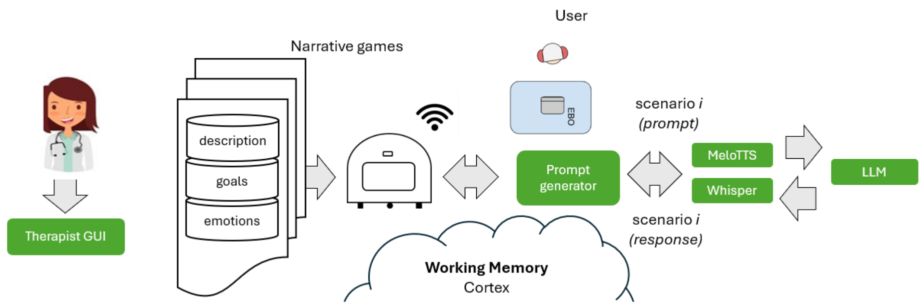

Fig. 1. Overview of the proposal described in this paper.

The EBO-based system presented in this paper represents a significant advancement in integrating these technologies for cognitive therapy. EBO can deliver highly personalized and adaptive cognitive therapies by incorporating the CORTEX cognitive architecture and utilizing LLMs like Llama and advanced ASR and TTS systems. The "therapist-in-the-loop" methodology ensures that the therapy is tailored to the specific needs of each user, enhancing its effectiveness and user satisfaction. This approach addresses traditional methods' limitations and sets a new standard for deploying social robots in cognitive therapy. Integrating these technologies allows EBO to provide a more dynamic and engaging therapeutic experience, potentially improving older adults' cognitive and emotional well-being with mild to moderate cognitive impairments [2]. Furthermore, AI-driven personalization in therapeutic programs, such as physical exercise for elderly patients with cancer, has shown significant benefits, highlighting the importance of tailored interventions [8].

3 Self-adaptive Cognitive Therapies with EBO Robot

The system presented in this paper integrates new human-robot interaction capabilities into the Cortex architecture that autonomously adapts therapy to older adults. This section provides a detailed overview of the system's architecture, functionalities, and AI techniques to enhance user engagement and therapeutic effectiveness. Fig. 1 shows a general schematic of the proposed system. First, the occupational therapist selects the type of game that will be developed in the therapy session. Next, the initial game prompt is generated and played through the TTS using the personalized data for each user. At this point, the iteration between the user and the robot begins and continues until the end of the session. During this process, MeloTTS, Whisper, and Llama capture data and update the robot's knowledge through the Cortex architecture. Once the game is completed, the therapist receives feedback from the session.

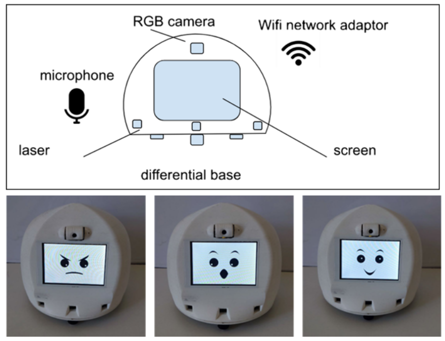

Fig. 2. The EBO social robot and some of the facial expressions.

3.1 System Architecture

The proposal is based on the Cortex cognitive architecture [5]. This architecture is based on a shared Working Memory (WM) representing the EBO robot's knowledge about the world. Different software agents access this WM to update or use this data to infer new knowledge. Thus, the main components of our system include the EBO robot hardware, the CORTEX cognitive architecture, and several AI-based agents for speech recognition, natural language processing, and adaptive interaction.

a) EBO robotic platform. EBO is a social robot designed in RoboLab, University of Extremadura. It is equipped with sensors, cameras, and actuators to interact with users dynamically and responsively. Its design incorporates expressive features to enhance social interaction and engagement. This ability to express emotions has proven highly beneficial in cognitive therapy with older adults, making a significant difference when our system is implemented in a robot, leading to more engaging interactions [10]. In this same study, we found that it has been very well received, having been tested with real users who provided positive evaluations and expressed a strong desire to engage with it.

Regarding the hardware, a co-creation methodology was employed to design the external form, resulting in a plastic casing with a diameter of less than 15 cm and a weight of under one kilogram to optimize user acceptance. The EBO robot's components include:

- A Raspberry Pi 3B+ that manages the control system for other hardware elements.
- A speaker and microphone for user interaction.
- A camera to capture visual data.
- A display screen to show images, such as emotional expressions.
- Laser sensors that provide distance measurements from surrounding objects.

b) Cortex architecture: Integrating AI Agents for Enhanced Adaptive Interaction. The CORTEX cognitive architecture offers a structured approach for designing, modularizing, and representing robotic activities and data [5]. Within this architecture, agents are tasked with specific roles and collaborate by utilizing a Working Memory. This WM holds the robot's knowledge, which can be predefined, derived from sensor inputs, or generated through agent actions. The WM is organized as a directed graph, where vertices represent metric or symbolic information and edges denote geometric or logical relationships. In this graph, vertices (or nodes) symbolize ontology concepts, while edges define their connections.

Using the CORTEX architecture, agents can share information and coordinate their actions to facilitate self-adaptive human-robot interaction during therapy sessions. Fig. 3 illustrates the CORTEX instance employed in this study. This working memory holds data from various agents, each contributing to the system's overall functionality and interaction capabilities. The agents include:

- Automatic Speech Recognition (ASR) Agent: This new agent accurately captures and interprets spoken language from users, which enables natural and fluid interactions between the user and our robot. In our proposal, the ASR agent uses Whisper, developed by OpenAI [9]. Whisper is an ASR system trained on 680,000 hours of multilingual and multitasking supervised data collected from the web.
- Text-to-Speech (TTS) Agent: The TTS Module synthesizes natural-sounding speech, enabling the robot to communicate effectively and empathetically with users. This capability is achieved by incorporating the MeloTTS technology [14]. MeloTTS, developed by MyShell.ai, is a text-to-speech library that converts written text into natural, fluent speech across several languages. Its primary feature is the generation of highly natural and expressive voices using deep learning techniques.
  The MeloTTS system operates through four main stages: first, the input text is processed to extract linguistic structures. Next, intermediate acoustic features representing the speech are generated. These features are then used to synthesize the final audio waveform. This architecture enables real-time synthesis on CPUs without requiring GPUs, ensuring efficient and high-quality speech output.
- Dialogue manager Agent: This agent, powered by Large Language Models (LLMs), generates meaningful and contextually appropriate responses during interactions. This enhances the robot's ability to engage users in therapeutic conversations and activities. In our proposal, we used a Llama developed by Meta [12] as our LLM and tested it with different models to adjust the conversations and games developed.
- Player and Game Data Integration (therapist-in-the-loop): This agent includes data relating to the older adult (player) and the selected game base, all through a user application specifically designed for the occupational therapist. This data ensures that the game interactions are tailored to the specific user, considering their cognitive abilities and personal preferences.

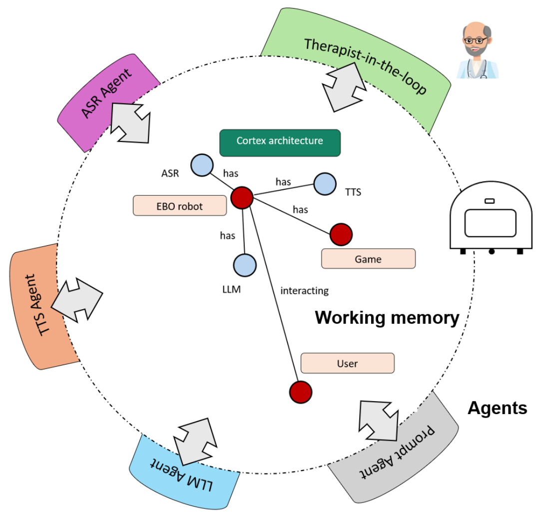

Fig. 3. The CORTEX architecture and the set of agents used in this proposal.

- Prompt Generator Agent: The working memory's data allows for the generation of personalized dialogs through the Prompt Generator Agent. This agent synthesizes the information from all sources (including the environment and the user's state) to create prompts for the LLM Agent that guide the interaction, ensuring it is engaging and suitable for the user's cognitive level.

These agents must be interconnected precisely, where CORTEX plays a crucial role. By reading real-time information from the graph, CORTEX facilitates this interconnection. Each agent will have a node with different attributes that change as the system operates. Other agents will detect changes in these attributes to determine when to act. The Working Memory is described in Fig. 4, where nodes and their associated attributes are detailed.

Interconnection of Agents to Ensure System Functionality The interconnection of the agents is designed to facilitate seamless communication and operation within the AI module system. Table 1 outlines the processes involved in this interconnection, illustrated in Fig. 5.

This structured interconnection ensures continuous and efficient agent interaction, enabling the system to function smoothly throughout the game.

4 Interactive Cognitive Games and Therapies: therapist-in-the-loop

The system features cognitive games designed to enhance memory, attention, and problem-solving skills, through interactive storytelling, task memorization,

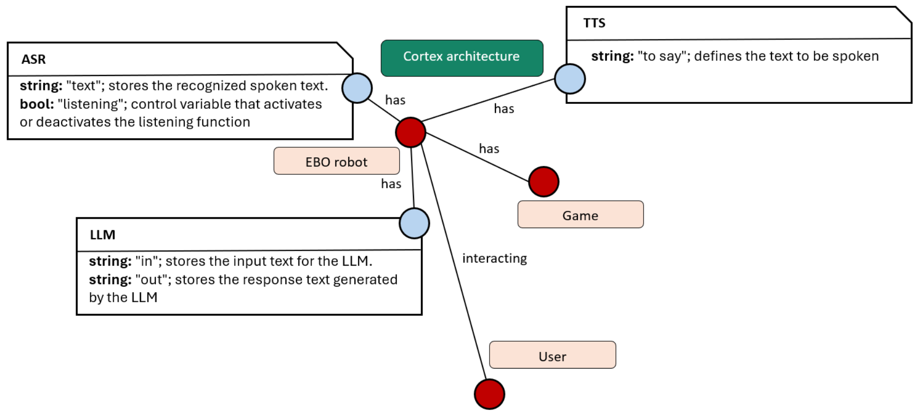

Fig. 4. The Working Memory defined in our use case. The CORTEX architecture implements three main nodes for interconnecting the ASR, TTS, and LLM agents.

and puzzles. Tailored to individual cognitive levels and interests, these games offer personalized and engaging experiences. Occupational therapists play a key role by inputting user data, selecting games, adjusting difficulty, and providing feedback. They can also alter the game flow via an interface connected to the Working Memory using a designated agent.

This paper details the implementation of three games focusing on verbal training for Activities of Daily Living (ADLs), which are critical for maintaining and enhancing the independence of older adults. ADLs refer to the essential daily tasks necessary for individuals to live autonomously and fulfill basic needs.

Game 1: ADLs Sequences This activity aims for the participant to correctly sort the steps of a basic ADL that EBO will present in an unsorted sequence. EBO will begin by announcing the activity and then randomly providing the steps. The participant's task is to verbally sort these steps correctly, ensuring they reflect the proper sequence for the activity. EBO will evaluate the answer and provide feedback to determine if the response is correct. Table 2 presents the initial prompt and the commencement of the game.

Game 2: True or False of ADLs The game aims for the participant to identify whether the statements related to basic daily life activities provided by EBO are true or false, reinforcing their knowledge and understanding of these activities. EBO will present a series of statements about various basic daily life activities, and the participant must determine if each statement is true or false. This exercise tests their comprehension and helps them retain critical information about daily living tasks. Each game will consist of five questions, with the correct answers recorded for further analysis and feedback. Table 3 shows this game's initial prompt and the first user's interactions.

Process	Description
Initialization and Response Generation	– The LLM agent receives the initial prompt from the Game agent and stores it in the LLM 'in' attribute. – The LLM agent then generates a response to initiate the game, storing this response in the LLM 'out' attribute.
TTS Processing	– The TTS agent monitors the LLM 'out' attribute. Upon detecting a change, the TTS 'to say' attribute is updated with the new response. – The TTS agent is programmed to add the string from TTS 'to say' to the playlist upon detecting this update. – Once the message finishes playing, the TTS agent sets the ASR 'listening' attribute to True.
ASR Processing	– The ASR agent begins listening when it detects that the ASR 'listening' attribute is activated. – After capturing the spoken input and detecting silence, the ASR agent updates the ASR 'text' attribute with the transcribed string.
Loop and Continuation	– The LLM agent re-enters the process by detecting the change in the ASR 'text' attribute. – The LLM agent updates the LLM 'in' attribute with the newly transcribed text from the ASR agent. – The system then reverts to the beginning of the loop, where the LLM agent generates a new response based on the updated prompt, continuing this cycle until the game concludes.

Table 1. Overview of the EBO Robot's Interactive System Processes

Game 3: The Shopping The game aims for the participant to identify the item that should not be present from a list of ingredients required to achieve a specific goal. Once correctly identified, they must calculate the total cost of the purchase in euros. This aids in improving planning, money management, and mathematical skills. EBO will start by announcing the meal to be prepared, and the participant will select the unnecessary item. Once all ingredients are chosen, EBO provides their prices in euros. With this information, the participant calculates the total cost of the purchase. Upon completion, EBO assesses the list and the calculation and offers its opinion to the participant. Table 4 shows the initial question and the start of this game

5 Experimental results and discussion

5.1 Comparison of the different LLM models evaluated

In this subsection, we will analyze the different LLM models we have tested, with the help of professional occupational therapists working in the group, to determine which one best meets our needs. Our evaluation will focus on each model's ability to generate engaging and effective games. By comparing their performance, we aim to identify and implement the model that best enhances our system and user experience.

To analyze and compare the different models studied, we focused on asking them to generate a single game: the sequences of daily life activities. Specifically,

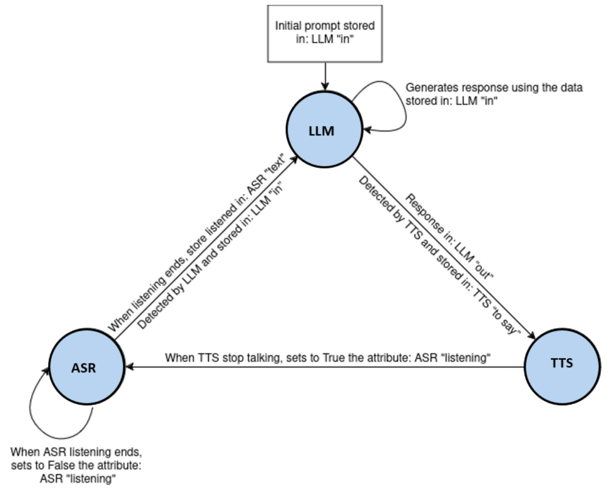

Fig. 5. Schematic showing the interconnection of the AI agents

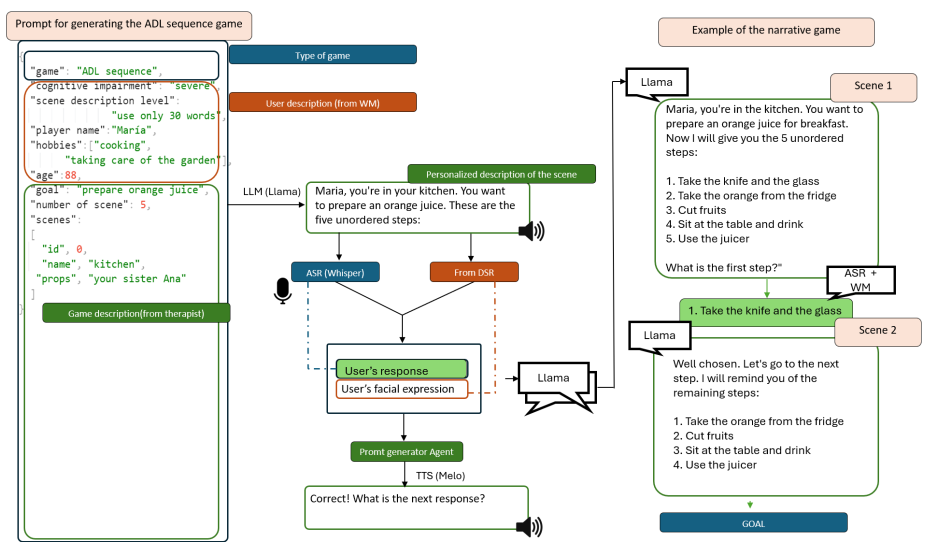

Fig. 6. The figure illustrates the cognitive therapy process with the EBO robot, showing the initial prompt of the game-generating agent along with the development of the game that leverages ASR/TTS or working memory.

Initialization prompt	I want you to play with me. You must present a scrambled sequence of steps to reach the goal that I indicate. I'll respond one by one. You tell me if it's correct or not. If it's correct, we continue. If it's wrong, we start over. I'll provide key data in JSON format. Customize the game based on that data. Don't tell me anything now. Just start the game
LLM response	Welcome to the game, Maria! Here's your first scrambled sequence of steps to prepare orange juice. Add sugar to taste. Cut the oranges in half. Wash the oranges. Squeeze the juice from the oranges. Pour the juice into a glass. **What's your first move?"**
First user intervention	Add sugar
Game continuation	Sorry, that's not the correct next step. Let's start over. Here's the scrambled sequence again: ...

Table 2. Initial prompt and fist interactions of the game designed for the ADLs sequences

Initialization prompt	I want you to play with me. You have to present me with a question related to a daily life activity that gets me to the goal I present. I'll respond true or false. You tell me if it's correct or not. If correct, we continue. If wrong, we start over. I'll provide you with the most important data in JSON format. Personalize the game according to that data. Don't tell me anything now. Just start the game.
LLM response	María. Let's start the "ADL True or False" game by preparing orange juice in the kitchen with your sister Ana! **Question 1: First, do you need to select ripe oranges from the fridge? True or false?**
First user intervention	True
Game continuation	Step 2: Great! Let's move on. Then, do you wash the oranges under the faucet? True or False?

Table 3. Table presenting the game designed for the True or False game of ADLs

we asked the LLM models to generate a 5-step sequence, and we tested this with three different models. The model parameters have been carefully selected to ensure the most appropriate and accurate outputs possible.

The three models under study are:

Llama 3

It'll be used Meta-Llama-3-8B-Instruct, a fine-tuned version specialized for instructional text generation, that will be executed on one of the two A100 80GB GPUs available in a rack server. The following key parameters are configured:

– *temperature=0.6*: is applied during text generation to control the randomness and ensure conservative outputs.
– *top_p=0.9*: Top-p probability is used for sampling, allowing for diverse outputs while maintaining coherence.
– *max_gen_len=None*: refers to the maximum generation length, meaning there is no explicit limit on the length of generated sequences.

The *tokenizer.model* is used to tokenize the input text into a format that the model can understand. It contains the vocabulary and rules necessary for converting raw text into numerical inputs that the model can process.

Initialization prompt	I want you to play with me. You have to present me with a shopping list to achieve the goal I indicate, with one false item. I have to find the false item. You tell me if it's correct or not. If correct, we continue, and you give me the prices in euros so I can calculate the total. If wrong, we start over. I'll provide you with the most important data in JSON format. Customize the game according to those data. Don't tell me anything now. Just start the game.
LLM response	Welcome, María! We're about to start our shopping game. Your goal is to prepare orange juice. Here's your first shopping list: Oranges Sugar Ice Lettuce Lemons **Can you identify the false item?"**
First user intervention	Lemons
Game continuation	Correct! Lemons don't belong in orange juice. Well done! Now, let me provide you with the prices in euros for the items on your list: Oranges - €2.50 Sugar - €1.20 Ice - €0.80 Lettuce - €1.00 **Please calculate the total.**

Table 4. Table presenting the game designed for The Shopping game

ChatQA-1.5-8B-Q8_0.gguf

It has been downloaded from Huggingface repository and loaded using LlamaCpp. During the conversation, a prompt template defines the structure of interactions, enabling seamless communication between the user and the AI assistant.

The model is triggered with user responses to generate coherent and relevant answers. As part of the setup, a specified embeddings model is initialized, responsible for converting text inputs into numerical representations. Additionally, the conversation history is managed using Chroma DB, which stores text embeddings for efficient retrieval and context maintenance. This testing has been conducted on a local computer. The system is equipped with 32.0 GiB of RAM, an Intel® Core™ i7-6800K CPU operating at 3.40 GHz with 12 cores, and an NVIDIA GeForce GTX 1080 Ti graphics card.

GPT-3.5

The last one, not directly implemented in the system, was GPT-3.5. The same prompt was tested in this model to compare the performance of the Llama models with this one. The file size of GPT-3.5 is estimated to be around 350 GB. This LLM is capable of understanding and generating text coherently and naturally. The obtained results are analyzed in the Table 5. The therapeutic aspects analyzed in this study include adaptability, quality of feedback, personalization of dialogue, continuity of discourse, and data recording.

Based on these results, we have decided to conduct tests using GPT-3.5 to assess user acceptance. Positive user acceptance would confirm the effectiveness of our approach, justifying further efforts towards the autonomous integration of large language models (LLMs) into the system. Future autonomous implementation may utilize Llama A100, which has also yielded promising results, although its limited memory capacity has currently prevented its deployment.

5.2 Evaluation of the System with Real Users

We tested our complete system with 10 volunteers from our research group using a Likert-scale questionnaire to evaluate various aspects of functionality and per-

Model	Response Time (seconds)	Dialogue analysis
Llama 3	14.3	This model fully adapts to the participant's level of cognitive impairment by tailoring the dialogue accordingly, but its effectiveness could not be correctly assessed due to a lack of memory. Despite this limitation, it demonstrates significant personalization by addressing the participant by name and gender. However, the absence of memory results in no continuity of discourse. It features good data recording capabilities, enabling subsequent evaluation of the therapy administered.
ChatQA-1.5-8B-Q8_0.gguf	33.2	This model does not adapt to the participant's level of cognitive impairment, resulting in poor feedback and easily lost conversation threads. It shows minimal personalization, as it does not address the participant by name. While there is short-term continuity of discourse, the system fails to function properly after three to four messages. Despite these issues, it maintains good data recording capabilities for the subsequent evaluation of the therapy provided.
GPT-3.5	5.1	This model fully adapts to the participant's level of cognitive impairment, generating a natural and dynamic conversation with good feedback. It demonstrates a high degree of personalization by addressing the participant by name and gender. The model maintains complete continuity of discourse, even recalling information from past interactions. Additionally, it features good data recording capabilities, allowing for effective subsequent evaluation of the therapy conducted.

Table 5. Comparison of the dialogues generated by each model

formance. Each participant played a selection of the previously described games with EBO, utilizing all the agents in the architecture. Following the gameplay, the survey included questions about user experience, system responsiveness, and overall satisfaction, providing a comprehensive assessment of the system's functionality and effectiveness. The summary of the results from the user satisfaction survey is presented in the Table 6

Question	Average value in the response (standard deviation)
How well did you feel in the interaction with EBO?	4.2 (0.63)
Did you find EBO's instructions and explanations clear and easy to understand?	4.7 (0.48)
Do you feel that the feedback provided by EBO during the game is adequate?	4.1 (0.99)
How easy was it to understand and answer EBO's questions during the game?	4.7 (0.48)
Did you find the questions appropriate for you?	4.7 (0.67)
Do you think EBO's language was appropriate?	4.3 (0.82)
Do you think the interaction with EBO could improve your skills and functions?	4.3 (0.67)
Would you recommend EBO activities and games to others?	4.8 (0.42)
Did you feel that EBO understood all your answers? (Speech recognition)	3.9 (1.1)

Table 6. Results of the interaction with EBO

The evaluation results of the interaction with EBO indicate a predominantly positive experience among participants. Most questions received high scores, particularly those related to the clarity of instructions and ease of understanding, averaging 4.7 with a low standard deviation of 0.48. This reflects strong consistency and overall clarity in EBO's communication. Additionally, the willingness to recommend EBO activities scored the highest, with an average of 4.8 and a standard deviation of 0.42, indicating high user satisfaction. However, the perception of EBO's understanding of responses was lower, with a score of 3.9 and a higher standard deviation of 1.1, suggesting variability and potential issues with voice recognition, especially for male voices. Overall, while EBO interactions are effective and positive, improvements in voice recognition are needed for a consistently satisfying user experience.

6 Conclusion

In this article, we have delineated the development of a system employing social robots like EBO for interactive cognitive games and therapies bolstered by Large Language Models (LLMs), signifying a promising pathway to address the cognitive exigencies of the elderly populace. By harnessing narrative generation, user response capture, and narration capabilities, personalized interactions are attained and tailored to individual preferences and cognitive capacities. The "therapist-in-loop" approach ensures a collaborative framework, where therapists contribute to initial configurations and autonomous therapy sessions are conducted, culminating in comprehensive summaries for subsequent analysis.

This personalized therapeutic approach seeks to heighten user experience and optimize the therapeutic efficacy of social robots in fostering cognitive and emotional well-being among older adults. The pilot study conducted with three different games underscores the potential effectiveness of this approach and indicates a significant stride toward refining cognitive therapies through innovative technological integration. Thus, the integration of LLMs within social robotics holds promise in advancing therapeutic interventions for older adults, representing a pivotal step toward addressing the escalating needs of this demographic group.

Acknowledgment

This work has been partially funded by the PDC2022-133597-C41 project and TED2021-131739-C22 R+D+i projects, supported by Spanish MICIU/AEI/10.13039/501100011033 and the European Union's "NextGenerationEU"/PRTR, by the PID2022-137344OB-C31 R+D+i project, supported by Spanish MICIU/AEI/10.13039/501100011033 "FEDER Una manera de hacer Europa", "FEDER/UE", and by the FEDER Project 0124_EUROAGE_MAS_4_E (2021-2027 POCTEP Program).

References

1. Barata, A.N.: Social robots as a complementary therapy in chronic, progressive diseases. Advances in Experimental Medicine and Biology pp. 105–120 (2019). https://doi.org/10.1007/978-3-030-24230-5_5

2. Cifuentes, C.A., Pinto, M.J., Céspedes, N., Munera, M.: Social robots in therapy and care. Preprint (2020). https://doi.org/10.1007/S43154-020-00009-2
3. Demange, M., Lenoir, H., Pino, M., Cantegreil-Kallen, I., Rigaud, A.S., Cristancho-Lacroix, V.: Improving well-being in patients with major neurodegenerative disorders: Differential efficacy of brief social robot-based intervention for 3 neuropsychiatric profiles. Clinical Interventions in Aging pp. 1281–1290 (2018). https://doi.org/10.2147/CIA.S152561
4. Dinesen, B., Hansen, H.K., Grønborg, G.B., Dyrvig, A.K., Leisted, S.D., Stenstrup, H., Schacksen, C.S., Oestergaard, C.U.: 'hi robot!'–use of a social robot for elderly with dementia: An exploratory study. Preprint (2022). https://doi.org/10.2196/preprints.36505
5. García García, J.C., Núñez Trujillo, P.M., Bachiller Burgos, P., Bustos García, P.: Towards the design of efficient and versatile cognitive robotic architecture based on distributed, low-latency working memory. In: Proceedings of IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC 2022). Santa Maria da Faira, Portugal (2022)
6. Kalik, J.: Effectiveness of cognitive stimulation personalized by the preexisting cognitive level in older adults. Topics in Geriatric Rehabilitation pp. 78–90 (2022). https://doi.org/10.1097/tgr.0000000000000345
7. Mahmoudi Asl, A., Molinari Ulate, M., Martin, M.F., van der Roest, H.G.: Methodologies used to study the feasibility, usability, efficacy, and effectiveness of social robots for elderly adults: Scoping review. Journal of Medical Internet Research **24** (2022). https://doi.org/10.2196/37434
8. Mendibil Esquisabel, B., Larrad, A., Arrieta, H., Zeberio Etxetxipia, I.: Onkofrail: Personalization of a physical exercise program in older patients with lymphoma. Hematological Oncology pp. 45–50 (2023). https://doi.org/10.1002/hon.3166_ot33
9. Radford, A., Kim, J.W., Xu, T., Brockman, G., McLeavey, C., Sutskever, I.: Robust speech recognition via large-scale weak supervision (2022)
10. Rodríguez-Domínguez, M., Bazago-Dómine, M., Jiménez-Palomares, M., et al.: Interaction assessment of a social-care robot in day center patients with mild to moderate cognitive impairment: A pilot study. International Journal of Social Robotics **16**, 513–528 (2024). https://doi.org/10.1007/s12369-024-01106-4, https://doi.org/10.1007/s12369-024-01106-4
11. Sumioka, H., Torresen, J., Shiomi, M., Chen, L.K., Nakazawa, A.: Editorial: Interaction in robot-assistive elderly care. Frontiers in Robotics and AI **9** (2022). https://doi.org/10.3389/frobt.2022.1020103
12. Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., Lample, G.: Llama: Open and efficient foundation language models (2023)
13. Tulsulkar, G., Mishra, N., Thalmann, N.M., Lim, H.E., Lee, M.P., Cheng, S.K.: Can a humanoid social robot stimulate the interactivity of cognitively impaired elderly? a thorough study based on computer vision methods. The Visual Computer pp. 1–12 (2021). https://doi.org/10.1007/S00371-021-02242-Y
14. Zhao, W., Yu, X., Qin, Z.: Melotts: High-quality multi-lingual multi-accent text-to-speech (2023), https://github.com/myshell-ai/MeloTTS
15. Zheng, Z.K., Zhu, J., Fan, J., Sarkar, N.: Design and system validation of rassle: A novel active socially assistive robot for elderly with dementia. Robot and Human Interactive Communication (RO-MAN) pp. 121–126 (2018). https://doi.org/10.1109/ROMAN.2018.8525819