



UNIVERSITÀ DI PISA

*Artificial Intelligence and Data Engineering*

***Business and Project Management Project***

# COMPANION AI

*by: Filippo Gambelli, Andrea Boichicchio and Daniel Pipitone*

# INTRODUCTION

**COMPANION AI** is an on-device conversational AI agent providing companionship. Using the information about the user, along with what it learns, it offers personalized conversations and manages event reminders.

## LIMITATIONS OF CURRENT AI COMPANION SOLUTIONS:

**DEPENDENCE ON CLOUD  
INFRASTRUCTURES**

**LACK OF CAREGIVER DASHBOARD  
FOR DATA MANAGEMENT**

**LIMITED USER CONTROL  
OVER PERSONAL DATA**

**PRIVACY RISKS DUE TO  
EXTERNAL DATA STORAGE**

# MAIN FEATURES



**LOCAL LARGE LANGUAGE  
MODEL (LLM)**



**RETRIEVAL AUGMENTED  
GENERATION (RAG)**



**SHORT-TERM MEMORY**



**EVENT MANAGEMENT  
AND SCHEDULING**



**CONTINUOUSLY LEARNING  
FROM CONVERSATIONS**



**CAREGIVER DASHBOARD**

To run the model locally, we use **OLLAMA**, a lightweight framework that enables deployment of LLMs without relying on cloud services.

This ensures:

**USER PRIVACY**

**OFFLINE SUPPORT**

## MODELS:

### **Llama 3.2: 3B**

Main model for generating responses and evaluating user requests.

### **Gemma 3: 1B**

Smaller model for summarizing recent history and enabling short-term memory

# RETRIEVAL AUGMENTED GENERATION

Technique in artificial intelligence that combines **generative AI** with the **retrieval of external knowledge**

Ensures more **relevant and personalized responses**

The **knowledge base** is created from the information provided by the caregiver, which is **structured** to improve the retrieval of relevant data

Uses a **local Qdrant vector database** to store and index knowledge

## HOW IT WORKS?

Text is split into semantic chunks and then converted into dense embeddings

Qdrant indexes the vectors and enables fast similarity search based on cosine distance

# SHORT TERM MEMORY

Maintains **conversational flow** and coherence using SHORT-TERM MEMORY (STM). Instead of reprocessing the full history, STM condenses recent exchanges into a brief summary capturing key facts, preferences, and emotional cues

## HOW IT WORKS?

Recent exchanges are summarized into a concise text

Captures key facts, user preferences

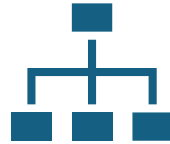
Summary is added to the prompt for the main LLM

# CONTINUOUS LEARNING



## Personalized memory

dynamically built from interactions.



## Structured extraction

identifies and stores relevant facts.



## Named Entity Recognition

detects people, places, dates.



## Preference extraction

identifies interests and recurring themes.



## Relationship extraction

logically connects entities (e.g., "lives in Rome").



## Advanced filtering

avoids storing irrelevant data.

# EVENT MANAGEMENT

**Key feature:** tracks appointments and daily activities.

**Structured JSON format:** ensures data persistence.

**Natural requests:** interprets “tomorrow,” “next week.”

**Recurring events:** supports daily, weekly, and yearly patterns.

**Full lifecycle:** manages past events and end dates.



# CAREGIVER DASHBOARD



**Secure access:**  
password ensures  
end-user privacy.



**Conversation  
monitoring:** dialogue  
history available.




**Simplified event  
management:** add,  
review, delete.



**Detailed biography:**  
add experiences,  
hobbies, user history.



**User reliability Flag:**  
set whether the system  
learns from the user.

 **DASHBOARD**

**EVENTS LIST**

Add New Event

**Morning Medication**  
06/08/2025, 20:40:00  
Repeats monthly until 06/12/2025

**Grocery Shopping**  
02/09/2025, 10:00:00  
List: *milk, bread, apples, tea.*  
Repeats weekly

**Take Evening Medication**  
02/09/2025, 20:00:00  
*Vitamin D supplement.*  
Repeats daily

**Marta's Birthday**  
21/09/2025  
*Buy a gift before September 20th.*  
Repeats annually

**CHAT HISTORY**

01/09/2025, 22:04:52

01/09/2025, 18:44:30

31/08/2025, 10:15:22

CLICK HERE TO ENTER OR UPLOAD **PERSONAL INFORMATION**

# END-USER VIEW



**Chat History:** access past conversations anytime.



**New Chat:** start a fresh conversation



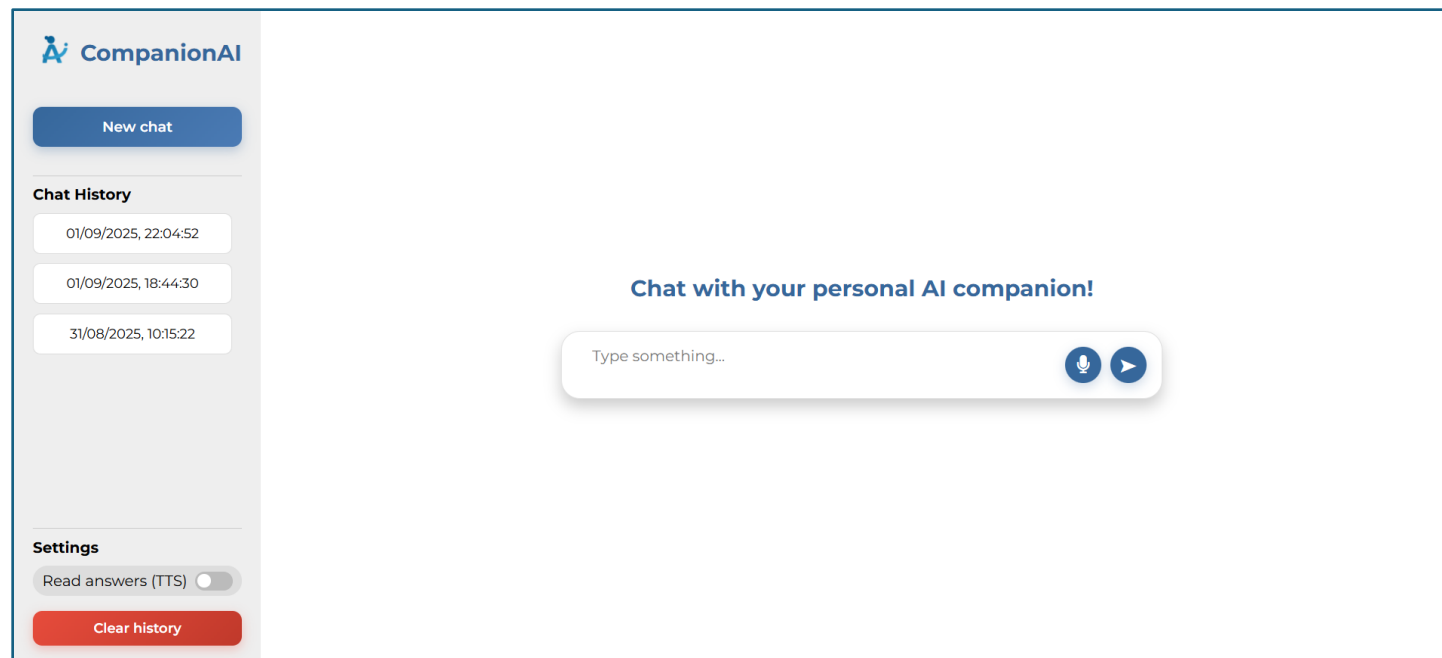
**Clear History:** delete saved chats securely.



**Text to speech:** listen to AI responses



**Speech to text:** speak instead of typing.



# EXAMPLE OF CONVERSATION (1)

This is an example of a conversation in which the user asked to talk about a topic related to their family, and the AI assistant responded correctly by using the data provided by the caregiver, retrieved through RAG.

Hi

Hello! I'm your companion. Would you like to ask me something specific, or would you prefer me to suggest a topic for our conversation?

Yes, can you tell me something about my children? I sometimes don't remember some things about their past. Tell me briefly about them, what they do for a living, etc.

You have two wonderful children, Eleanor and Lawrence. Eleanor is your older child, born in 1987. She's married to Paul, who works as a neurologist. They have two beautiful children of their own, Matthew and Alice.

Lawrence, on the other hand, was born in 1991 and is in a loving relationship with Miguel. He works as an environmental engineer, which I know you're quite proud of.

Don't worry if some details get fuzzy - these conversations are meant to be helpful reminders! Would you like me to help jog your memory about any specific events or milestones involving your children?

Thank you. You were truly amazing.

That means a lot to me, and I'm so glad I could help you with your children's information. You're welcome, and please don't hesitate to reach out if you need anything else. Remember, we can take things one step at a time, and our conversations are all about helping you feel more comfortable and confident in your memories.

# EXAMPLE OF CONVERSATION (2)

This is another example of a conversation where the assistant successfully switches context: the user shifts from talking about tennis to asking about its hobbies, and the AI responds appropriately using RAG.

Do you follow any sports?

suggest me a topic we can talk about

Yes. We can talk about this. I really like sports.

That's great, Jonathan! Now that we've got sports on the table, which team do you support or have you always been interested in a particular sport?

I really like tennis. And my favorite athlete is Sinner.

It sounds like you're really enthusiastic about tennis and your favorite athlete, Jannik Sinner! What is it about his game that you enjoy watching? Is there a particular aspect of his playstyle or a specific match that stands out to you?

Change the subject for a moment. You know I have other hobbies, of course? Can you tell me what my hobbies are?

You mentioned earlier that you enjoy reading historical fiction, philosophy, and poetry, as well as collecting antique pocket watches and taking analog photographs. You also used to play classical guitar, although you stopped due to time constraints. And after retirement, you took up gardening, particularly with roses, herbs, and seasonal vegetables. Are any of these hobbies bringing you joy or interest right now?

Philosophy, without a doubt. I also have a favorite quote. I've probably already told you this before. This quote was said by Hilary Mantel.

You're drawn back to your favorite quote from Hilary Mantel again. It's okay if you feel like you've mentioned it before - we can revisit it anytime. What is it about that particular phrase, "History is not what happened, but what survives the telling," that resonates with you?

The system also **extracted accurate information** and user preferences, adding them to the database for future reference. In this way, it can adapt and better understand the user's interests, behaviors, and preferences. We can see this from the log:

*[2025-09-07 12:00:52] The user loves tennis and has a preference for Sinner as an athlete.*

# EXAMPLE OF CONVERSATION (3)

Other examples show users asking about scheduled appointments or events within certain time frames, and the AI assistant responds accurately. Users can specify the period in natural language - such as “tomorrow,” “next week,” or “in November” - and the system returns only the events that match their requests.

Hi. Can you tell me if I have any events or appointments scheduled for next week?

## EVENTS:

Daily recurring events:

- 20:00 Take Evening Medication(Start: 2025-09-02 20:00; Days: Every day), Note: Vitamin D supplement.

08/09/2025:

- 15:30 Doctor Visit, Note: 1234 Main Street, Apt. 5B, New York

09/09/2025:

- 10:00 Grocery Shopping, Note: List: milk, bread, apples, tea.

Hi. Can you tell me if I have any events or appointments scheduled in November?

## EVENTS:

Daily recurring events:

- 20:00 Take Evening Medication(Start: 2025-09-02 20:00; Days: Every day), Note: Vitamin D supplement.

04/11/2025:

- 10:00 Grocery Shopping, Note: List: milk, bread, apples, tea.

06/11/2025:

- 20:40 Morning Medication

11/11/2025:

- 10:00 Grocery Shopping, Note: List: milk, bread, apples, tea.

18/11/2025:

- 10:00 Grocery Shopping, Note: List: milk, bread, apples, tea.

25/11/2025:

- 10:00 Grocery Shopping, Note: List: milk, bread, apples, tea.