University of Salerno

Department of Information and Electrical Engineering and Applied Mathematics

Master's Degree in Computer Engineering

Enterprise Software Architectures Curriculum

GRADUATION THESIS

Knowledge Graphs for Time Assistive Management in Cognitive Impairment

Supervisor

Prof. Sabrina Senatore

Candidate

Alexandra Tarantino

0622701904

Academic Year 2023-2024

Abstract

DESCRIPTION OF THE PROBLEM

Mild Cognitive Impairment (MCI) represents an intermediate stage between normal cognitive aging and dementia, characterized by a decline in memory and executive functions that, although they do not completely impair daily autonomy, significantly affect quality of life. Among the most critical difficulties are deficits in time management, manifested through an altered perception of the duration of activities, difficulties in planning and organizing daily routines, and erroneous monitoring of the passage of time, resulting in forgetfulness of commitments, failure to follow therapies, and a persistent state of disorientation.

Assistive technologies (ATs), particularly Time Assistive Technologies (TATs) as electronic calendars, voice reminders, and virtual assistants, although useful for remembering commitments and deadlines, often have unintuitive interfaces and limited customization, forcing the caregiver to frequent manual interventions to update content and adapt it to the user's needs.

There emerges a need for technological solutions that favor natural interaction, accessible even to those with MCI: simple conversational systems that facilitate the entry, consultation, and modification of commitments. Such solutions should make time management more intuitive and autonomous, reducing caregiver burden and improving the daily well-being of people with MCI.

FRAMING OF THE THESIS IN THE CONTEMPORARY TECHNICAL SCENARIO

Large Language Models (LLMs) have revolutionized natural language processing in recent years, but they suffer from "hallucinations" and lack of transparency about sources, compromising the reliability of answers. The Retrieval-Augmented

Generation (RAG) paradigm has overcome these limitations by integrating a mechanism for retrieving information from unstructured external documents by semantic search, however, it ignores logical relationships among information, providing fragmented and unrelated contexts.

GraphRAG, an evolution of RAG proposed in 2024, combines vector search with traversal of a Knowledge Graph, overcoming the constraints of textual documents through a semantic map in which nodes represent entities and relationships capture conceptual dependencies. This approach improves both accuracy and completeness while enabling source traceability and explainability of the generation process.

The present work is among solutions that leverage LLMs for natural language processing, but aims to overcome their limitations in terms of accuracy and transparency. Central to the proposal is the adoption of the GraphRAG paradigm, based on the use of a Knowledge Graph built with the support of a dedicated ontology. The user with MCI is offered a conversational Time Assistive Technology, accessible via natural language, which allows personal information to be stored and guarantees complete, reliable and personalized answers based on the individual profile.

PERSONAL CONTRIBUTION OF THE CANDIDATE

The main contribution of this thesis is the design and implementation of an original GraphRAG-based solution for temporal assistance of patients with Mild Cognitive Impairment, a recent approach not yet adopted in the literature.

The developed system consists of two components: the Knowledge Graph Construction Component, which transforms natural language input into a personal knowledge graph, and the GraphRAG Component, which implements the GraphRAG approach to retrieve information and answer user questions.

A distinctive element is the integration of semantic support in the populating phase: a domain ontology (TAMOntology) has been defined and implemented to guide the extraction, classification, and insertion of data into the graph, ensuring semantic consistency and facilitating inference. Through the use of the ontology,

the information retrieval phase is also optimized, improving the accuracy, completeness, and relevance of the answers provided to the user.

This approach proves to be more reliable than LLM-only based solutions, which are prone to "hallucinations," and is more efficient and comprehensive than traditional RAGs that do not exploit relational links between data, providing people with MCI with natural language interaction, correct, personalized answers, and with source transparency.

DESCRIPTION OF THE APPLICATION/EXPERIMENTAL CONTENTS OF THE THESIS WORK

The developed system consists of a set of integrated components built in the Python environment and based on the use of the Neo4j database for knowledge graph management. A complete pipeline for graph construction and querying has been implemented. The process is based on a specially defined domain ontology, the TAMOntology, which provides the semantic schema underlying the knowledge representation. A prototype web user interface, built with Streamlit, has also been constructed, which allows interactive visualization of the implemented features.

To evaluate the benefits of using a knowledge graph in a RAG approach, an experimental protocol was implemented that, through the RAGAS framework, computes specific metrics for RAG-type systems: Context Precision, Context Recall, Answer Relevancy, Faithfulness, and Semantic Similarity.
The experimentation also involved the creation of ten synthetic profiles, each consisting of thirty interactions, employed to generate the knowledge graphs in Neo4j. Ten test datasets, each consisting of ten questions, were constructed with which the systems could be compared and metrics calculated.