2023 IEEE International Conference on High Performance Computing & Communications, Data Science & Systems, Smart City & Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys)

Empowering Personalized Health Data Queries with Knowledge Graph and GPT-Enhanced Voice Assistant

Xin Wang
Department of Computer Science
North Dakota State University
Fargo, USA
xin.wang.2@ndsu.edu

Wordh Ul Hasan
Department of Computer Science
North Dakota State University
Fargo, USA
wordh.hasan@ndsu.edu

Juan Li
Department of Computer Science
North Dakota State University
Fargo, USA
j.li@ndsu.edu

Rasha Hendawi
Department of Computer Science
North Dakota State University
Fargo, USA
rasha.hendawi@ndsu.edu

Yang Du
Department of Computer Systems and Software Engineering
Valley City State University
Valley City, USA
yang.du@vcsu.edu

Bo Xie
School of Nursing
The University of Texas at Austin
Austin, USA
boxie@utexas.edu

Cui Tao
School of Biomedical Informatics
University of Texas Health Science Center at Houston
Houston, USA
cui.tao@uth.tmc.edu

***Abstract*— The proliferation of digital health technologies has led to an abundance of personal health data. However, querying and retrieving specific health-related information from disparate sources can be challenging and inconvenient, particularly for older individuals unfamiliar with technology. While ChatGPT offers a conversational interface, it lacks domain-specific knowledge, including personalized health information. To address this limitation, we present a novel approach that combines a knowledge graph and GPT to enable personalized health queries. Our solution utilizes a personal knowledge graph as a comprehensive knowledge source and fine-tunes GPT to provide accurate responses. We have implemented a voice assistant mobile app incorporating this knowledge graph-assisted GPT model and conducted initial feasibility testing.**

***Keywords*— personal health data, knowledge graph, Generative Pre-trained Transformer, graph database, chatbot**

I. Introduction

In the current era of digital health and wearable devices, individuals have unprecedented access to a vast array of personal health data [1], [2]. For instance, health apps generate data such as daily activity levels, heart rate, sleep patterns, and nutrition intake. Wearable devices, like fitness trackers and smartwatches, track physical activity, exercise routines, and vital signs. Medical records store comprehensive health histories, diagnoses, and treatment plans. Lab data, including blood tests and imaging results, provide insights into individuals' physiological conditions. Additionally, insurance data contains information on coverage, claims, and reimbursement records. Today, approximately 30% of the world's data volume is being generated by the healthcare industry [3]. By 2025, the compound annual growth rate of data for healthcare will reach 36%. That's 6% faster than manufacturing, 10% faster than financial services, and 11% faster than media & entertainment [3].

Despite the availability of these data sources, the fragmented nature of the data across various platforms presents a significant challenge for individuals seeking a holistic view of their health. For example, consider a user who wants to assess their lifestyle and make improvements based on recommendations from their primary doctor. This user would typically need to search for their Electronic Health Records (EHR) data to retrieve the doctor's suggestions. They would then have to navigate through various health apps and wearable devices to access data on their physical activity, sleep patterns, and nutrition. Finally, they would need to synthesize this information to make informed lifestyle choices. This fragmented process can be time-consuming and cumbersome, often requiring manual efforts and technical proficiency.

The primary objective of this research is to develop a knowledge graph-assisted GPT model that enables individuals to query and retrieve personalized health information effortlessly. The proposed system aims to bridge the gap between users and their health data by leveraging the power of natural language processing and knowledge graph technology. By integrating a personal knowledge graph and fine-tuning the GPT model, we strive to provide accurate and tailored responses to users' health-related queries, regardless of their technological proficiency.

ChatGPT has gained significant popularity across various domains, revolutionizing the way people interact with conversational AI. From customer support to content generation, ChatGPT has showcased its capabilities in understanding and generating human-like responses. While ChatGPT excels in generating responses based on general

This work was supported by the U.S. National Science Foundation (NSF) with award numbers: 2218046.

979-8-3503-3001-4/23/$31.00 ©2023 IEEE
DOI 10.1109/HPCC-DSS-SmartCity-DependSys60770.2023.00149

knowledge, it falls short when it comes to addressing personalized questions that require access to personal private data. ChatGPT cannot tap into an individual's specific health information. Furthermore, there have been instances where ChatGPT produces plausible-sounding but incorrect or nonsensical answers. This phenomenon, often referred to as "hallucination," poses significant risks when it comes to critical domains such as medical advice or accurate health-related facts. To ensure the reliability and safety of using AI in such contexts, it is crucial to address these limitations and enhance the accuracy of responses.

By integrating various data sources, such as health apps, wearable devices, electronic health records (EHRs), lab data, and insurance records, into a comprehensive knowledge graph, we create a unified representation of an individual's health information. This knowledge graph serves as a contextualized and structured knowledge foundation that can be used to fine-tune GPT. During the fine-tuning process, we train GPT using a specialized dataset derived from the integrated knowledge graph. This dataset consists of personalized health data, ensuring that the model learns to generate accurate and personalized answers. By incorporating the knowledge graph during the fine-tuning phase, we enhance the model's understanding of health-related concepts, relationships, and context, enabling it to provide context-aware and reliable responses.

By combining the strengths of GPT's natural language processing capabilities with the contextual knowledge from the personal health knowledge graph, we aim to create a chatbot system that can bridge the gap between general knowledge and personalized health information. This approach ensures that users receive accurate, reliable, and tailored responses to their health queries while maintaining privacy and data security.

In the following sections, we will delve into the details of constructing the personal health knowledge graph, fine-tuning GPT using the knowledge graph, and conducting feasibility testing to evaluate the effectiveness and usability of our approach.

II. Related Work

Knowledge graph-based approaches have gained attention in the healthcare domain for organizing and querying personal health information.

[4] introduces and explores the use of knowledge graphs for question answering in the context of electronic health records (EHR). The authors propose a graph-based approach for EHR question answering, which leverages the natural representation of relationships in knowledge graphs to facilitate more accurate and intuitive querying compared to table-based approaches. The authors compare the performance of EHR question answering model on both the table-based and graph-based datasets. The experimental results demonstrate that the graph-based approach significantly improves accuracy compared to the table-based approach, even without modifications to the model architectures. Another work [5] tries to address the U.S. opioid epidemic using a knowledge graph-based approach. The approach leverages the Opioid Drug Knowledge Graph (ODKG) as a comprehensive network encompassing opioid-related drugs, active ingredients, formulations, combinations, and brand names. They use the ODKG to normalize drug strings in a clinical data warehouse and showcase the use of ODKG to generate summary statistics of opioid prescription trends across US regions.

[6] proposed a blockchain-based personal health knowledge graph to address the challenges associated with managing and integrating personal health data (PHD) from various sources. The approach utilizes knowledge graphs to structure and integrate different types of PHD, and they used blockchain to ensure data privacy and security. The proposed approach aims to provide a comprehensive view of an individual's health by integrating and managing diverse PHD sources. [7] presents a medical knowledge-based dialogue system that acts as a health assistant chatbot. The chatbot analyzes users' reported disease symptoms and provides personalized medical advice and nutrition suggestions. By utilizing a medical knowledge graph, the chatbot can ask relevant questions to narrow down the search range of possible diseases. The system efficiently matches symptoms against its medical knowledge base and recommends suitable diagnoses. Similarly, [8] introduces KnowHealth, a knowledge graph-based question-answer platform designed to manage health data for elderly people. KnowHealth utilizes an ontology to define entities and relations in the domain of aged diseases. Health-related information is crawled and extracted from various sources, and entities and relations are used to construct a knowledge graph. The platform provides a historical behavior-driven question-answering system that accurately retrieves and reasons answers by analyzing and extending the intention of questions.

ChatGPT, a language model developed by OpenAI, has gained significant attention in the healthcare domain due to its potential applications in improving patient care, medical research, and healthcare operations [9]. [10] examines the advantages and limitations of using ChatGPT in health education. The study utilizes expert panel discussions and literature review to assess the benefits and drawbacks of ChatGPT in healthcare educational domains. The findings indicate potential benefits such as personalized learning, improved clinical reasoning, and enhanced understanding of complex concepts in medical education. However, the limitations identified across all healthcare disciplines include data privacy concerns, the risk of biased or inaccurate content generation, and potential negative effects on critical thinking and communication skills among students.

Sharma et al. [11] explore the use of ChatGPT in providing healthcare services to mariners in the maritime industry. The authors highlight how ChatGPT can revolutionize maritime healthcare by offering personalized and prompt healthcare services. By leveraging the expertise and conversational capabilities of ChatGPT, virtual consultations with healthcare professionals can be enabled, allowing for the analysis of health data. The integration of ChatGPT technology in maritime healthcare has the potential to transform the way seafarers receive care and support. However, the study acknowledges the need to address challenges associated with implementing ChatGPT-powered healthcare services in the maritime sector. Another research [12] developed a voice assistant app powered by GPT to empower caregivers of individuals with Alzheimer's

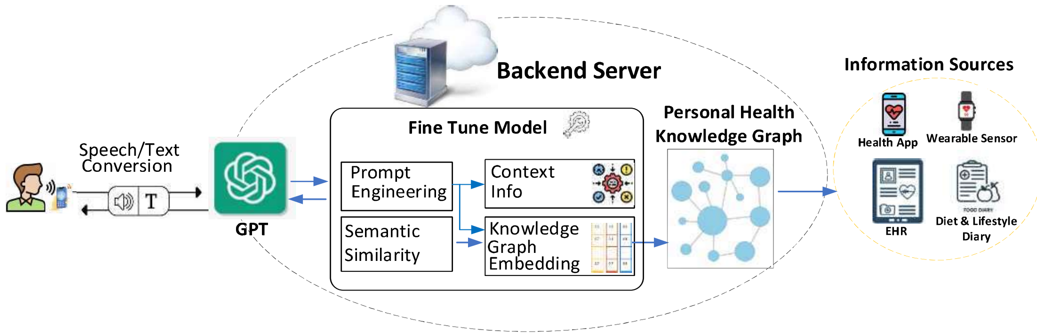


Fig. 1. System architecture

Disease and Related Dementias (ADRD). The proposed voice assistant aims to facilitate caregivers of ADRD in accessing shared experiences and practical tips from peers, providing them with valuable insights. Initial evaluation of the app has shown promising results, indicating its feasibility and potential impact on caregivers.

III. System Design

In our system, depicted in Fig. 1, users can effortlessly access their health knowledge through a voice assistant. By speaking their queries, the voice assistant converts the speech to text and forwards it to GPT for processing. GPT, utilizing the user's query, consults the knowledge base comprising the user's personalized health data to provide relevant answers. To build the knowledge base, we construct a comprehensive knowledge graph that connects various health-related data through nodes and edges. To optimize GPT's utilization of the knowledge graph, we convert it into a knowledge graph embedding. This embedding serves as a contextual reference for GPT during the query processing, enabling a more effective response.

A. Construction of Personal Health Knowledge Graph

We start with real-time data collection from diverse sources, including health apps, wearable devices, medical records, lab data, and insurance records. These sources provide valuable information about the patient's health condition and context, enabling a holistic view of their well-being. To ensure consistency and interoperability, the raw data from different sources undergoes transformation into a unified format based on a common ontology – HealthOnto.

As shown in Figure 2, this ontology defines relevant health concepts, such as symptoms, diseases, treatments, medications, etc., along with their relationships. We reused HealthOnto that was designed in our previous research [6]. This ontology serves as the foundation for the personal health knowledge graph ensuring a uniform framework for organizing and displaying personal health data. To design the ontology, we followed the HL7 FHIR [13] standard. This approach enhances consistency and interoperability across diverse sources of personal health data. The HL7 FHIR-based ontology proposed in [14] serves as our foundation. We extended the HL7 FHIR ontology to include additional aspects of an individual's health, such as profiles, lifestyle interventions, healthcare providers, and data generated by health-related wearable devices. For example, the expanded ontology now includes classes like Physical Activity, Diet, Smoking Status, and Alcohol Consumption, along with their sub-classes and properties, which provide valuable information about an individual's lifestyle.

Based on HealthOnto we can align and link the data from various sources for building a coherent and comprehensive knowledge graph. Semantic similarity and advanced entity resolution techniques [15], [16] are employed to match and connect similar or related data points. This step addresses challenges arising from variations in naming conventions or identifiers across different sources.

Figure 3 illustrates a segment of a personalized health knowledge graph utilized to capture individual health insights. In this example, the graph pertains to a user named Heinz, encompassing both specific and confidential data, such as allergies and family medical history. Furthermore, due to his type 2 diabetes diagnosis, the knowledge graph comprehensively stores details about Heinz's primary physician, upcoming appointments, lab results, prescribed medications, and more. Heinz takes a proactive approach to monitor his well-being and enhance his lifestyle by leveraging wearable devices,

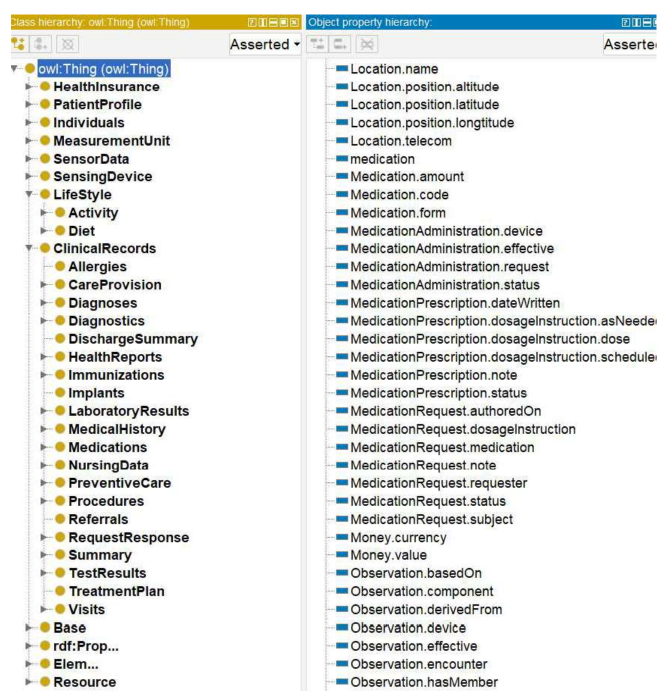


Fig. 2. Part of the HealthOnto ontology

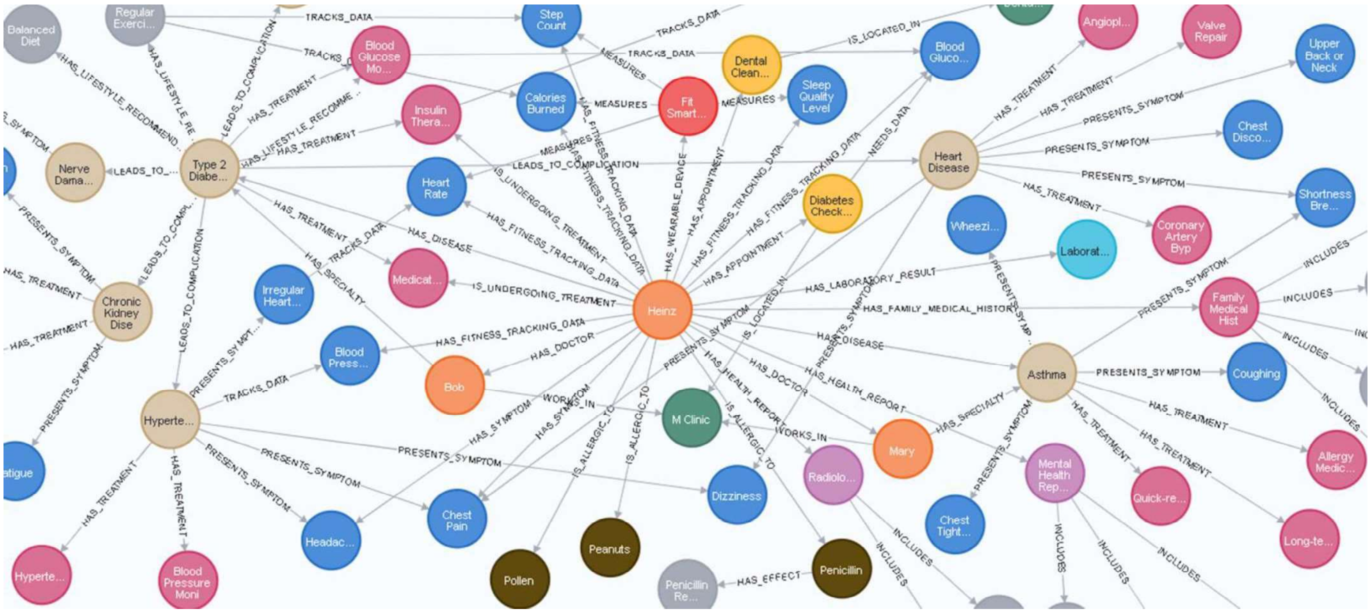


Fig. 3. Part of the knowledge graph with personal health data

health applications, and similar tools to track vital metrics like heart rate, blood glucose levels, step count, and sleep quality. Additionally, the knowledge graph encompasses pertinent data regarding the conditions Heinz is managing, encompassing facets such as type 2 diabetes complications and symptoms associated with asthma. As Heinz engages with inquiries concerning his health status, treatments, appointments, and related subjects, the system delivers precise responses anchored in his private information. This level of tailored insight goes beyond the capabilities of a general GPT model, ensuring accurate and personalized assistance for Heinz's distinct health journey.

We augment the personalized health knowledge graph (PHKG) with external knowledge, leveraging resources such as the Systematized Nomenclature of Medicine (SNOMED) [17], [18], a widely used clinical terminology standard. This enriches the graph with additional information, including general facts about diseases or treatments, contextual details like weather or air quality, and other relevant insights.

B. *Multi-Modal Data Storage*

For real-time sensor data and other dynamic information, we adopt a high-performance and scalable data storage solution. This ensures rapid and efficient access to the constantly evolving health data, which is crucial for providing up-to-date information through the voice assistant. Dynamic data refers to real-time or near-real-time data that is continuously generated and updated, such as sensor data from wearable devices, health app inputs, vital signs, activity levels, and other dynamic health-related metrics. To effectively store and manage this constantly evolving data, we employ a dynamic data storage solution that is designed to handle high volumes of incoming data and support rapid data retrieval. Specifically, a time-series database is employed for storing dynamic health data as it is specifically designed to handle time-stamped data points. Each data point is associated with a timestamp, allowing for easy retrieval and analysis of data over time. Time-series databases are optimized for efficient storage and retrieval of time-series data, making them well-suited for sensor data and other time-sensitive health metrics. By implementing a dynamic data storage, our voice assistant mobile application can efficiently capture, store, and retrieve real-time health data, providing users with accurate and timely health information through the power of voice interactions.

On the other hand, static data, such as the patient's medical history and past records, are stored in a secure and reliable data storage repository. This static data is essential for building a comprehensive understanding of the patient's health profile, which aids in delivering accurate and personalized responses. By combining these data storage approaches, our voice assistant mobile application effectively utilizes multi-modal access to health data.

C. *Health Knowledge Graph Embedding*

Creating a knowledge graph embedding is a crucial step in accurately capturing the semantics and relationships between concepts and entities in the health domain. Knowledge graph embedding techniques aim to represent entities and relationships from a knowledge graph in a continuous vector space, where the spatial proximity of embeddings reflects the semantic similarity and relationships between them.

In the proposed system, we adopt one popular approach, Sentence-BERT [19], a transformer model trained by SNLI corpus [20], TriviaQA [21], wikiHow [22], etc. Sentence-BERT fine-tunes BERT in a Siamese/Triplet network architecture [19] which is a specialized neural network structure designed for learning similarity and dissimilarity between two pieces of data like sentences. Based on the different downstream tasks, the Sentence-BERT used different datasets, downstream structures, and loss functions to produce semantically meaningful sentence embeddings.

Our knowledge graph embedding aims to quantify how relevant a knowledge sentence relates to a user's question.

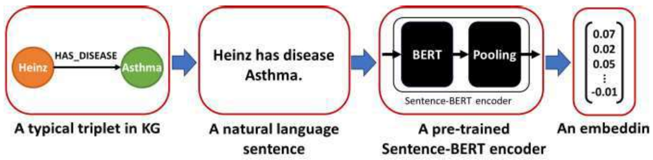


Fig. 4. The process of covering knowledge graph triples into embeddings

Knowledge graphs consist of a large number of interconnected triplets (a head entity such as "Heinz", a relationship such as "HAS_DISEASE", and a tail entity such as "Asthma"). The knowledge graph embedding technique translates each triplet into a semantic statement in the form of a subject-predicate-object sentence (e.g., "Heinz has disease asthma"), and further converts the sentence into a dense vector as shown in Figure 4. By measuring the similarities (spatial proximities) between the user-question vector and all the knowledge vectors, the sentences semantically relevant to the user's question can be found. The sentences with the top n largest similarities will be sent as context to the GPT model via prompt engineering.

D. *Tuning the GPT Model*

In our pursuit of providing natural language responses to user queries, we leverage the power of the GPT model. To address the inherent challenges associated with GPT, such as the inability to offer personalized information and the potential to generate plausible sounding yet incorrect answers, we employ a set of strategies. Central to our approach are two critical components: prompt engineering and the integration of contextual embeddings. These elements collectively contribute to elevating the accuracy and relevance of the model's responses. Our prompt engineering strategy revolves around a fundamental principle: instructing GPT to generate responses that are as truthful as possible, hinging solely on the context and the content encapsulated within the knowledge graph, presented in the form of embeddings.

Consider the scenario where a user raises a query: "What was my heart rate this afternoon? Was it normal?" To guide the GPT model in generating precise, coherent, and satisfactory responses, the prompt must strike a balance between offering pertinent details and curbing the tendency of over-generating, which could result in erroneous answers. Our prompt's design takes a strategic approach. It initially presents a comprehensive enumeration of specific information in the "Context" segment. This information comprises highly individualized details that are outside the realm of the GPT model's knowledge. This tailored domain-specific context acts as a pivotal foundation. Simultaneously, the prompt underlines the significance of logical reasoning, harnessing the GPT model's capacity to make sound inferences based on the provided factual data. This directive enhances the model's capability to address intricate queries effectively within the designated domain. However, to preclude the model from offering misguided responses that often stem from an inclination to respond even when information is absent, our prompt employs a precautionary measure. It explicitly instructs the GPT model to "answer the question as truthfully as possible," while also incorporating a fail-safe option of "I don't know" when an answer cannot be identified.

By imparting this guidance, we ensure that GPT derives its responses from the amalgamation of contextually relevant information contained within the knowledge graph embeddings. This approach enhances the coherence and veracity of the generated responses, effectively minimizing the possibility of unfounded or speculative answers.

IV. System Evaluation

To gauge the effectiveness of our approach, we developed a functional prototype in the shape of a mobile application. Rigorous evaluation procedures were conducted, encompassing a multitude of use case scenarios as well as comprehensive user studies.

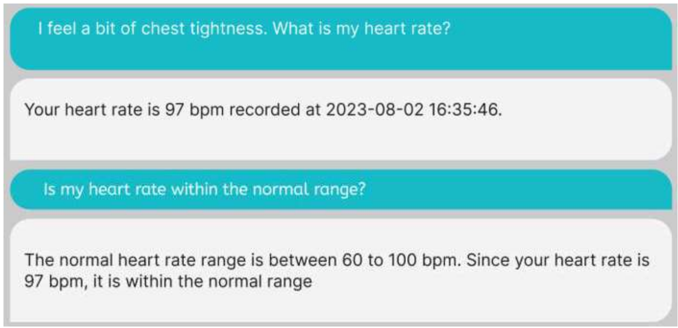
A. *Prototype Impementation*

Our implementation involves creating databases, back-end service, and a user-friendly mobile chatbot application, synergistically designed to deliver a seamless and intuitive user experience. We used Neo4j [23] to store the personal health knowledge graph and employed InfluxDB [24] to process all types of time series data. The back-end service developed using the Django web framework [25], acts as a crucial hub for data management and query processing. It interfaces between the user's personal health data and the GPT model. It ensures the integration and transformation of data from various sources into the knowledge graph, maintaining up-to-date health data for query processing. Additionally, it manages GPT model communication, training, and personalized response generation. For user interaction, we built a cross-platform mobile app using Flutter, Google's open-source UI toolkit. This chatbot interface enables users to submit natural language health queries. Flutter's flexibility, expressive UI, and native performance are ideal for swift development of natively compiled applications across multiple platforms. The app is accessible on both Android and iOS devices.

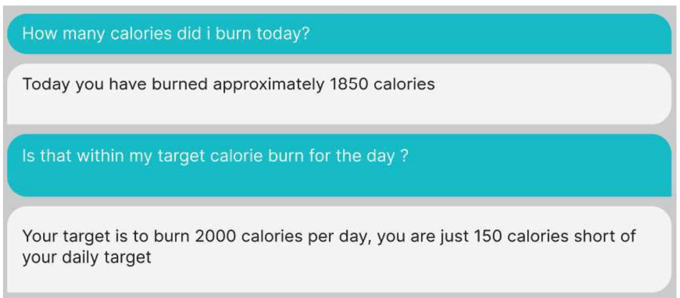
B. *Use Case Study*

We assessed our system's effectiveness through use case study based on various health data query scenarios. A hypothetical user, Heinz, is a 60-year-old male managing type 2 diabetes and hypertension. He employs a wearable device for monitoring blood glucose, heart rate, and activity levels. His health app tracks dietary habits and medication schedules, while his medical records reside in an electronic health record system. Using our system, Heinz's real-time wearable and health app data are integrated into a personal health knowledge graph. He interacts with our chatbot mobile app, seeking natural language responses regarding his health, medications, and more. The system leverages the knowledge graph and GPT model to furnish Heinz with personalized, accurate, and reliable answers.

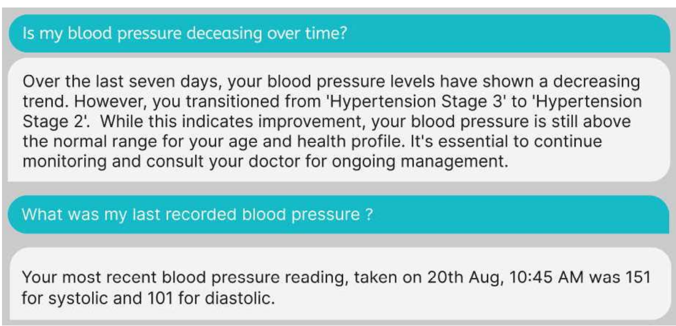
For instance, Heinz feels chest tightness and inquires about his recent heart rate. As shown in Fig.5 (a), the system accesses wearable data from the knowledge graph, contextualizing the response by indicating his heart rate falls within the standard range. This relieves Heinz from manually interpreting data or seeking medical advice. Similarly, Heinz inquires about his daily calorie burn. The system accesses fitness tracker data, offering the calorie count and comparing it to his daily goal, as shown in Fig. 5 (b). Such instant access and interpretation aid Heinz in meeting his health objectives. Further, as illustrated in



(a)



(b)



(c)

Fig. 5. Example use cases

Fig. 5 (c), Heinz wants to track his blood pressure trajectory. Drawing from historical data in the knowledge graph, the system informs him that his blood pressure has improved, moving from 'Hypertension Stage 3' to 'Hypertension Stage 2'. Despite improvement, values still exceed recommended limits for his age and profile. The system's concise yet insightful summaries assist Heinz in making informed health choices.

Through this use case study, we demonstrate our system's value in delivering personalized and contextually relevant health insights. This approach empowers users like Heinz, simplifying health management and informed decision-making.

C. Pilot Study

Before public deployment, particularly for systems involving personal health data, internal testing is vital to uncover and rectify potential issues. A pilot study with lab members is an initial and crucial step to assess system effectiveness, spot operational glitches, and gain valuable feedback for enhancements.

a) Comparative User Study:

To rigorously assess our system's efficacy and compare it with existing health information retrieval methods, we conducted a comprehensive pilot study involving lab members. This internal evaluation served as a crucial initial step in identifying potential issues and refining our system.

The study encompassed four distinct categories of health-related tasks, designed to mirror real-world scenarios. These categories covered Personal Health Profile, Lifestyle and Wellness, Medical Procedures and Treatments, and Clinical Records and Medical History. Participants performed these tasks using both our Virtual Assistant (VA) system and Traditional Methods (TM), allowing us to collect data on completion time and information accuracy. The traditional methods include browsing university and local hospital health portals, checking wearable devices apps like smartwatches, smartphone health applications, and other established means of interacting with health-related data. The results are listed in Table I.

Across all four categories, the VA consistently outperformed TM in terms of completion time, without compromising information accuracy. For example, in the Personal Health Profile category, the VA achieved a 100% completion rate with an average time of 13.5 seconds, while TM lagged at 41.5 seconds. A similar trend was observed in the Lifestyle and Wellness and Clinical Records and Medical History categories. The efficiency of the VA was particularly evident in the Medical Procedures and Treatments category, where it maintained a 100% success rate and accuracy, completing tasks in just 11.2 seconds on average. In contrast, TM's accuracy slightly dropped to 80% and required an average of 60.1 seconds.

These results underscore the VA's ability to swiftly retrieve accurate health data across diverse categories. By integrating various data sources into a unified platform, the VA minimizes the time required for information access, compared to TM where data might be dispersed across different tools and documents. However, it's important to note that this study's participants were tech-savvy graduate students, which may limit the generalizability of our findings. Our next steps include expanding the study to encompass a more diverse user group, ensuring robustness and applicability in various contexts.

TABLE I. COMPARATIVE PERFORMANCE OF VIRTUAL ASSISTANT (VA) SYSTEM AND TRADITIONAL METHODS (TM) IN PERSONAL HEALTH DATA RETRIEVAL

Scenario	# of tasks	Completion Rate (%)		Completion time (seconds)		Information Accuracy (%)	
		VA	TM	VA	TM	VA	TM
Personal health Profile	10	100	100	13.5	41.5	100	100
Lifestyle and wellness	10	100	100	10.6	51	100	100
Medical Procedures and treatments	10	100	100	11.2	60.1	100	80
Clinical Records and medical history.	10	100	100	12.4	44	100	100

b) User Experience Evaluation:

During the user experience testing phase of our pilot study, we thoroughly assessed the interaction aspects of our virtual assistant. For this evaluation, we employed established metrics like the User Experience Questionnaire (UEQ) [26] and the System Usability Scale (SUS) [27]. While the UEQ helps understand user sentiments about a system, the SUS measures system usability. To address chatbot-specific evaluations, the Chatbot Usability Questionnaire (CUQ) [28] was introduced, focusing on factors such as personality, onboarding, navigation, comprehension, replies, error resolution, and bot intelligence.

The CUQ, mirroring the SUS's structure, comprises 16 statements assessed on a five-point scale. Odd-numbered questions represent positive attributes, while even-numbered ones denote potential limitations. CUQ scores, calculated based on user responses, range from 0 to 100, with higher scores indicating better user experience. Our user studies' results, detailed in Table II, show consistently positive trends, with users finding the chatbot engaging, helpful, and easy to use. Some users noted occasional confusion. Overall, the chatbot's CUQ Score across users was high, with most ratings surpassing 87%, indicating a highly favorable user experience.

TABLE II. USER STUDIES RESULTS USING CHATBOT USABILITY QUESTIONNAIRE (CUQ).

	Questions \ Users	U1	U2	U 3	U4	U5
1	The chatbot's personality was realistic and engaging	4	5	4	5	4
2	The chatbot seemed too robotic	2	1	2	2	2
3	The chatbot was welcoming during initial setup	4	5	4	5	4
4	The chatbot seemed very unfriendly	2	2	1	2	1
5	The chatbot explained its scope and purpose well	5	5	5	4	5
6	The chatbot gave no indication as to its purpose	1	1	2	2	1
7	The chatbot was easy to interact and navigate	5	5	5	5	5
8	It would be easy to get confused when using the chatbot	1	3	2	2	2
9	The chatbot understood me well	4	4	5	3	4
10	The chatbot failed to recognize a lot of my inputs	2	2	2	1	1
11	Chatbot response were useful, appropriate and informative	5	5	5	5	5
12	Chatbot response were not relevant	1	2	1	1	1
13	The chatbot coped well with any errors or mistakes	4	4	5	4	4
14	The chatbot seemed unable to handle any errors	2	1	2	1	2
15	The chatbot was very easy to use	5	5	5	5	5
16	The chatbot was very complex.	1	2	1	2	1
CUQ Score for each user out of 100		87.5	87.5	89.1	85.9	89.1
Total CUQ Score		87.8±1.3				
Median Score		87.5				

V. CONCLUSIONS

The increasing availability of personal health data due to digital health technologies has created a challenge in efficiently querying and retrieving specific health-related information from various sources. This is particularly problematic for older individuals less familiar with technology. While ChatGPT offers a conversational interface, it lacks domain-specific knowledge, including personalized health information. To address this, we propose a unique solution that combines a knowledge graph and GPT to enable personalized health queries. Our approach uses a personal knowledge graph as a comprehensive knowledge source and fine-tunes GPT to provide accurate responses. A mobile app incorporating this system was developed and feasibility testing conducted. By integrating various health data sources into a knowledge graph and fine-tuning GPT, we enable context-aware and reliable responses. This bridges the gap between general knowledge and personalized health information, facilitating accurate and tailored answers for health-related queries.

REFERENCES

[1] J. Greiwe and S. M. Nyenhuis, "Wearable Technology and How This Can Be Implemented into Clinical Practice," *Curr Allergy Asthma Rep*, vol. 20, no. 8, p. 36, 2020, doi: 10.1007/s11882-020-00927-3.

[2] S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, "Big data in healthcare: management, analysis and future prospects," *J Big Data*, vol. 6, no. 1, p. 54, 2019, doi: 10.1186/s40537-019-0217-0.

[3] Ben Hirschler, "Healthcare's Data Tsunami." https://www.brunswickgroup.com/healthcare-data-i20729/ (accessed Feb. 22, 2022).

[4] J. Park, Y. Cho, H. Lee, J. Choo, and E. Choi, "Knowledge graph-based question answering with electronic health records," in *Machine Learning for Healthcare Conference*, PMLR, 2021, pp. 36–53.

[5] M. R. Kamdar *et al.*, "A knowledge graph-based approach for exploring the US opioid epidemic," *arXiv preprint arXiv:1905.11513*, 2019.

[6] J. Li, V. Pandey, and R. Hendawi, "A Blockchain-Based Personal Health Knowledge Graph for Secure Integrated Health Data Management," in *IEEE Conference on ICT Solutions for eHealth (ICTS4eHealth 2023)*, 2023.

[7] L. Bo, W. Luo, Z. Li, X. Yang, H. Zhang, and D. Zheng, "A knowledge graph based health assistant," in *Proceedings of the AI for Social Good Workshop at Neural IPS, Vancouver, BC, Canada*, 2019.

[8] C. Li, S. Hang, X. Hu, D. Chu, and H. Zheng, "Knowhealth: a knowledge graph based question-answer platform for elderly people," in *Mobimedia 2019: 12th EAI International Conference on Mobile Multimedia Communications, Mobimedia 2019, 29th-30th June 2019, Weihai, China*, European Alliance for Innovation, 2019, p. 154.

[9] M. Cascella, J. Montomoli, V. Bellini, and E. Bignami, "Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios," *J Med Syst*, vol. 47, no. 1, p. 33, 2023.

[10] M. Sallam, N. Salim, M. Barakat, and A. Al-Tammemi, "ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations," *Narra J*, vol. 3, no. 1, pp. e103–e103, 2023.

[11] M. Sharma and S. Sharma, "Transforming maritime health with ChatGPT-powered healthcare services for mariners," *Ann Biomed Eng*, vol. 51, no. 6, pp. 1123–1125, 2023.

[12] K. T. Zaman, W. U. Hasan, J. Li, and C. Tao, "Empowering Caregivers of Alzheimer's Disease and Related Dementias (ADRD) with a GPT-Powered Voice Assistant: Leveraging Peer Insights from Social Media," in *2023 IEEE Symposium on Computers and Communications (ISCC)*, 2023, pp. 1–7. doi: 10.1109/ISCC58397.2023.10218142.

[13] "HL7 FHIR Release 5." https://hl7.org/fhir/overview.html (accessed Mar. 25, 2023).

[14] V. Kilintzis, A. Kosvyra, N. Beredimas, P. Natsiavas, N. Maglaveras, and I. Chouvarda, "A sustainable HL7 FHIR based ontology for PHR data," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2019, pp. 5700–5703.

[15] X. Wang *et al.*, "Artificial Intelligence-Empowered Chatbot for Effective COVID-19 Information Delivery to Older Adults," *International Journal of E-Health and Medical Communications (IJEHMC)*, vol. 12, no. 6, pp. 1–18, 2021.

[16] R. Hendawi, S. Alian, and J. Li, "A smart mobile app to simplify medical documents and improve health literacy: System design and feasibility validation," *JMIR Form Res*, vol. 6, no. 4, p. e35069, 2022.

[17] "SNOMED." https://www.snomed.org/

[18] E. Hovenga, H. Grain, and T. Beale, "Chapter 4 - Fragmented global standards development organisations," in *Roadmap to Successful Digital Health Ecosystems*, E. Hovenga and H. Grain, Eds., Academic Press, 2022, pp. 65–96. doi: https://doi.org/10.1016/B978-0-12-823413-6.00025-2.

[19] N. Reimers and I. Gurevych, "Sentence-bert: Sentence embeddings using siamese bert-networks," *arXiv preprint arXiv:1908.10084*, 2019.

[20] S. R. Bowman, G. Angeli, C. Potts, and C. D. Manning, "A large annotated corpus for learning natural language inference," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Lisbon, Portugal: Association for Computational Linguistics, Sep. 2015, pp. 632–642. doi: 10.18653/v1/D15-1075.

[21] M. Joshi, E. Choi, D. S. Weld, and L. Zettlemoyer, "TriviaQA: A Large Scale Distantly Supervised Challenge Dataset for Reading Comprehension," *CoRR*, vol. abs/1705.03551, 2017, [Online]. Available: http://arxiv.org/abs/1705.03551

[22] M. Koupaee and W. Y. Wang, "Wikihow: A large scale text summarization dataset," *arXiv preprint arXiv:1810.09305*, 2018.

[23] J. Webber, "A Programmatic Introduction to Neo4j," in *Proceedings of the 3rd Annual Conference on Systems, Programming, and Applications: Software for Humanity*, in SPLASH '12. New York, NY, USA: Association for Computing Machinery, 2012, pp. 217–218. doi: 10.1145/2384716.2384777.

[24] S. N. Z. Naqvi, S. Yfantidou, and E. Zimányi, "Time series databases and influxdb," *Studienarbeit, Université Libre de Bruxelles*, vol. 12, 2017.

[25] D. Rubio, *Beginning Django*. Springer, 2017.

[26] B. Laugwitz, T. Held, and M. Schrepp, "Construction and evaluation of a user experience questionnaire," in *HCI and Usability for Education and Work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, November 20-21, 2008. Proceedings 4*, Springer, 2008, pp. 63–76.

[27] J. R. Lewis, "The system usability scale: past, present, and future," *Int J Hum Comput Interact*, vol. 34, no. 7, pp. 577–590, 2018.

[28] S. Holmes, A. Moorhead, R. Bond, H. Zheng, V. Coates, and M. Mctear, "Usability Testing of a Healthcare Chatbot: Can We Use Conventional Methods to Assess Conversational User Interfaces?," in *Proceedings of the 31st European Conference on Cognitive Ergonomics*, in ECCE '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 207–214. doi: 10.1145/3335082.3335094.