Journal of Technology and Informatics (JoTI)
Vol. 7, No. 2, October 2025
P-ISSN 2721-4842
E-ISSN 2686-6102

Semantic Knowledge Fusion in Healthcare: A Hybrid Approach for Connected Medicine

Blaise Muhala Luhepa[1*], John Bukasa Kakamba[2], Deo Munduku Munduku[3], Daniel Mazono Magubu[4], Albert Ntumba Nkongolo[5], Herman Matondo Mananga[6], Djonive Munene Asidi[7]

[1,4,5]Department of Mathematics, Statistics and Computer Science, University of Kinshasa, Kinshasa, DR Congo
[2]Endocrinology, Metabolism, and Nuclear Medicine, Kinshasa University Clinics, Kinshasa, DR Congo
[3'6]Department of Observatory of Morbidity phenomena,One Health Institute for Africa, DR Congo
[7]Department of Exploration and Production, Faculty of Oil, Gas and Renewable Energies, University of Kinshasa, DR Congo
blaise.muhala@unikin.ac.cd[1*], johnbukasa73@gmail.com[2], deomundukuobama@gmail.com[3], danielmazono@gmail.com[4], albertntumba@gmail.com[5] herman.matondo@unikin.ac.cd[6], djonive.munene@unikin.ac.cd[7]

Article Information

Article History:
Received : July 31th 2025
Revised : September 23th 2025
Accepted : October 27th 2025
Published : October 31th 2025

***Correspondence:**
blaise.muhala@unikin.ac.cd

Keywords:
Artificial Intelligence in Healthcare, Data Fusion, Medical Ontology, Tabular Data, Intelligent Medical Assistant

10.37802/joti.v7i2.1182

Journal of Technology and Informatics (JoTI)
P-ISSN 2721-4842
E-ISSN 2686-6102
https://e-journals.dinamika.ac.id/index.php/joti

***Abstract*:**

In a context where connected medicine requires increasingly explainable, accurate, and responsive systems, this paper presents an applied experimental research focusing on the development and evaluation of a hybrid intelligent assistant for healthcare data fusion. The study is based on the parallel combination of two data paradigms: classical tabular structures and their ontological equivalent. Using an intelligent assistant, we simultaneously query a medical dataset on diabetes in tabular form and the same dataset translated into an OWL ontology that can be queried using SPARQL. The aim is to demonstrate that the synchronised combination of these two models not only provides a more complete response but also one that is better contextualised and clinically exploitable. The research follows an experimental methodology, involving the implementation, testing, and comparative evaluation of both models on 300 questions classified by increasing complexity (simple, complex, and very complex). The results reveal a relevance rate above 99%, with an average response time suited to medical use. This work highlights the potential of hybrid architectures in connected health and paves the way for new decision-making assistants that fully exploit the semantic richness of medical knowledge.

INTRODUCTION

Diabetes represents one of the most pressing public health challenges of the 21st century. In 2023, an estimated 537 million adults worldwide had diabetes, a figure that is rising steadily, with projections exceeding 640 million by 2030 [1]. The majority of cases (over 90%) are of type 2 diabetes, which is largely associated with behavioural and environmental factors [2]. In sub-Saharan Africa, under-diagnosis rates exceed 60%, seriously compromising patients' quality of life [3]. Despite the emergence of digital tools, management remains hampered by the fragmentation of medical data and the absence of integrated intelligent systems capable of providing clinicians with effective assistance [4]. In this context, connected medicine opens up new prospects for personalised monitoring and clinical decision support [5]. However, this transformation requires tools capable of exploiting heterogeneous data tabular, textual, and semi-structured and of reasoning beyond simple factual extraction. Solutions based solely on the analysis of Excel files or SQL databases are limited in their ability to capture the semantic complexity inherent in medicine [6]. Conversely, symbolic approaches based on medical ontologies, such as SNOMED CT or OMOP, enable explicit reasoning, but remain difficult to use for clinicians who are not computer scientists [7]. Recent literature explores hybrid approaches, and models based on Knowledge Graphs have shown their effectiveness in federating and structuring complex clinical data [8].

Problem statement. Despite these advances, there is still a lack of integrated systems that can simultaneously exploit both tabular and ontological representations of medical knowledge in an explainable and clinically usable way. Current approaches either focus on data structure without semantic reasoning or rely on ontologies that remain too complex for real clinical environments. This fragmentation prevents the development of intelligent systems capable of combining speed, semantic depth, and interpretability.

Research gap. While the Federated Virtual Knowledge Graph (FVKG) paradigm [9,10] has demonstrated the technical feasibility of linking tabular and RDF data through mappings, there is little empirical evidence of hybrid systems capable of performing parallel reasoning and delivering explainable medical insights. No prior study has provided a local, operational implementation bridging tabular querying and ontological inference in a unified medical assistant.

Research questions and urgency. This work therefore addresses three key questions: (1) How can a hybrid assistant simultaneously query tabular and ontological data to generate accurate and clinically explainable results? (2) To what extent does semantic fusion improve the relevance and response time of medical queries compared to isolated models? (3) How can such a system contribute to building trustworthy, locally deployable intelligent assistants for connected healthcare? The urgency of this research lies in the growing need for interpretable and reliable AI in medical decision support, particularly in resource-limited settings where clinicians require transparent tools that combine speed and comprehension.

To address this gap, our study proposes a hybrid medical assistant capable of querying, in parallel, two sources derived from the same dataset: an Excel table representing the clinical data of diabetic patients and an OWL ontology representing the same dataset in a semantic model. The GPT engine merges both responses to produce a unified, contextualised, and explainable synthesis. The case study involves 100 patients with various forms of diabetes and 300 medical questions (100 simple, 100 complex, 100 very complex). The system achieves an overall accuracy of 99%, with average response times below six seconds.

This performance is supported by three complementary components: a tabular pipeline (via *Pandas*) ensuring rapid access to structured data, an ontological pipeline (via *SPARQL*) enabling complex semantic queries, and a GPT-based fusion engine that orchestrates the synthesis of both responses into a coherent medical interpretation. This model builds upon previous work demonstrating the benefits of paradigm fusion [11,12], leverages recent structures such as Personal Knowledge Graphs [13], and applies lightweight ontologies in a local context [14]. Finally, it is grounded in a rigorous evaluation framework derived from ontology engineering and medical artificial intelligence research [6]. By combining these technologies, our contribution shows that parallel, local, and explainable semantic fusion can not only optimise diabetes management but also establish the foundations for a new generation of intelligent, human-centred medical assistants.

STATE OF THE ART

Connected medicine uses artificial intelligence (AI) systems to interpret a variety of medical data, with the aim of optimising care and diagnosis. There are three main areas of research: models based on tabular data, ontologies and Knowledge Graphs, and hybrid architectures that combine these two paradigms. Structured records (CSV, EHR) fed by supervised models (decision trees, deep networks) show solid results for predicting complications associated with diabetes. However, their lack of capacity to provide semantically rich explanations limits their usefulness in clinical practice [15]. Ontologies (SNOMED CT, OMOP, LOINC) and Knowledge Graphs facilitate interoperability and allow expert reasoning. For example, the study by Spoladore et al (2024) in Artificial Intelligence in Medicine demonstrates the positive impact of ontology-based decision-making systems for nutritional and glycaemic monitoring of diabetic patients [16].

On the other hand, Botha et al (2024) explore through an in-depth review the impact of AI models on patient rights and safety, highlighting key ethical challenges [17]. Wang et al.'s (2023) work with MediTab strengthens tabular predictions through semantic enrichment [18]. SeFNet (Woźnica et al., 2023) introduces a technical framework for aligning tabular and semantic features to support machine learning ontology integration [19]. Complementarily, Afandi et al. (2024) highlight the interpretive dimension of data relationships using network analysis, reinforcing the role of relational structures in hybrid knowledge systems. The approach of Lepetit Ondo et al. 2025 enables transparent federation of tabular and RDF data, promoting confidentiality and consistency via SPARQL [20]. In the context of diabetes, Rad et al. (2024) integrate digital twins and patient-centric graphs, improving real-time blood glucose monitoring [21]. Finally, Qin et al (2025) have developed a diabetic Q&A system combining Neo4j and LLM, achieving over 85% accuracy for entity recognition and almost 89% for intention classification [22]. Challenges and limitations

- Semantic alignment: the robustness of SPARQL mappings remains a major challenge, particularly for complex alignments [6].
- Answer fusion: current hybrid systems are often limited to juxtaposition or weighting. The use of an API based on an LLM model such as GPT to generate a unified and explanatory synthesis remains largely unexplored.
- Clinical explicability: the construction of a traceable meta-analysis, which documents the origin and justification of responses, is still largely underexplored.

Table 1 below summarises these features, highlighting their main objectives, their respective strengths and their structural and functional limitations.

Table 1. Comparison of tabular, ontological and hybrid approaches in healthcare decision support systems

Model	Objectives	Forces	Limits
Tabular	Prediction, speed, simplicity	Reliable, well controlled	Not very semantic, limited explicability
Ontological	Inference, consistency, transparency	Formal reasoning, interoperability	Complexity, limited clinical adoption
Hybrid	Greater relevance & confidentiality	Synergy of the advantages of the two paradigms	Complex alignments, immature fusion

Our proposal is part of this dynamic, a medical assistant for diabetics, querying tabular and ontological sources in parallel, and merging the answers via an API based on LLM to offer contextualised, explained and clinically validated results.

METHODS

a) Type of research

This study is an applied experimental research combining elements of system design, semantic modelling, and empirical evaluation. It aims to build and assess a hybrid intelligent assistant for connected medicine through a real clinical use case on diabetes management. The approach follows a design–build–evaluate paradigm typical of computer science and information systems engineering.

b) Overview

Our methodology is based on the construction, implementation, and evaluation of a hybrid intelligent assistant capable of querying, in parallel, two types of data structures representing the same medical knowledge:
(1) a tabular dataset (Excel file extracted from an electronic medical record on diabetes), and
(2) an ontology derived from the same data source.
The main objective is to evaluate, in a comparative and combined way, the ability of each data representation to produce relevant, explainable, and rapid answers to clinical questions expressed in natural language.

c) Methodological steps

The proposed architecture is modular and hybrid, combining heterogeneous sources of medical data (structured files and ontologies) and semantic reasoning capabilities with a generative AI engine.

It is based on three major functional areas:

(1) Semantic knowledge centre: the knowledge base is modelled in the form of an ontology designed in Protégé, and used via Apache Jena for dynamic SPARQL queries. It is generated from a dataset of patients suffering from diabetes, taken from structured electronic records (EHR).

(2) Interface and interpretation centre: users (e.g. healthcare professionals) interact via a user interface (UI) connected to an API incorporating an LLM model. This API provides natural language processing, semantic reformulation of queries and explanatory merging of responses from the two sources.

(3) Tabular data centre: the initial formats (XLSX, CSV, XML, JSON) feed both the direct tabular structure and the ontology, guaranteeing the consistency of the content analysed.

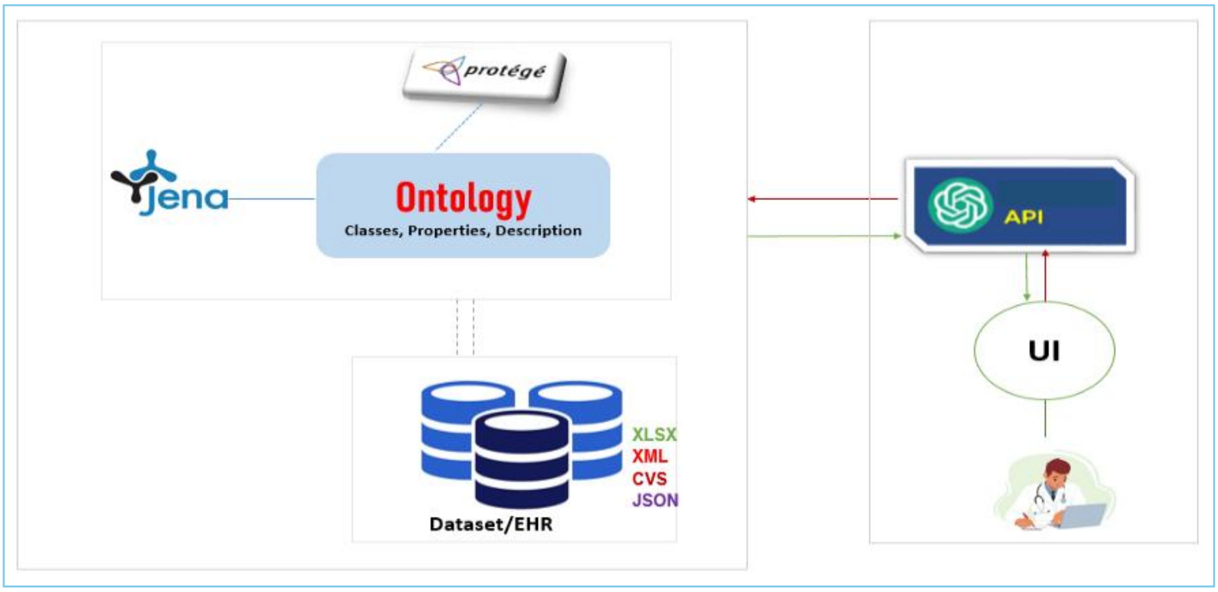

Figure 1. General architecture of the hybrid medical decision support system

1. Initiation phase: data collection and conceptual framework

The dataset used in this research was collected at the City Medical Laboratory in Kinshasa, and includes detailed information on diabetic patients, such as demographic attributes (age, sex, weight), medical indicators (blood glucose, complications, treatments), and diagnostic history. From a conceptual perspective, the framework relies on knowledge representation and semantic integration theories, aiming to bridge the gap between structured data and symbolic reasoning. The hypothesis underlying the experiment is that parallel querying of tabular and ontological representations improves interpretability and performance in medical question answering.

2. Development phase: architecture design

The proposed architecture is **modular and hybrid**, combining heterogeneous sources of medical data (structured files and ontologies) with semantic reasoning capabilities orchestrated by a generative AI engine. It is composed of three functional layers:

a) Semantic knowledge centre : The knowledge base is modelled as an OWL ontology designed in Protégé and hosted in Apache Jena for SPARQL queries. It is generated from the patient dataset derived from electronic health records (EHR).
b) Interface and interpretation centre : Users (clinicians, researchers) interact via a user interface (UI) linked to an API that integrates a Large Language Model (LLM). This API manages natural language understanding, reformulation, and semantic fusion of answers.
c) Tabular data centre : Original clinical data (Excel, CSV, JSON) are accessed directly through *Pandas* to ensure structural consistency between the relational and ontological sources.

3. Implementation phase

The hybrid assistant was developed in Python 3.10, using rdflib, pandas, openai, and matplotlib for processing and visualization. The ontology was created in Protégé, and semantic queries were executed via Fuseki Jena. The interface was implemented in Streamlit, enabling interactive input of medical questions and real-time display of responses.

The GPT-based fusion engine synchronises both data flows (tabular and semantic) to generate a unified, explainable answer. This design allows seamless interaction between local data and semantic reasoning without external dependencies.

4. Evaluation phase

To evaluate system performance, 300 medical questions were formulated and validated by an endocrinology expert. These questions were classified into three cognitive complexity levels (simple, complex, and very complex). Three main metrics were used for quantitative evaluation:

a) Response time (seconds) – measures latency between question and answer;
b) Relevance rate (%) – proportion of answers judged correct by medical assessors;
c) Response rate (%) – number of valid answers generated versus total questions asked.

Results showed an average response time of under six seconds and an accuracy rate of 99%, confirming the efficiency of semantic–tabular fusion for medical reasoning.

Ontological transformation

The transformation of the tabular set into an ontology took place in several phases:
(a) Identification of concepts (Patient, Treatment, Diabetes, Complication, etc.),
(b) Creation of classes and properties (e.g. aForTreatment, aDiabetes, aComplication),
(c) Automatic generation of RDF triplets from Excel rows.

How the assistant works

The intelligent assistant is structured into three parallel modules. Table 2 presents the three fundamental modules of the system: an NLP parser in charge of natural language understanding, a tabular engine for direct querying of structured files, and an ontology engine exploiting the inference capabilities of SPARQL on an OWL ontology. These modules work simultaneously to enrich the final response.

Table 2. Main functions of processing modules in hybrid architecture

Module	Main function
1. NLP Analyser	Natural language processing, extraction of key entities
2. Tabular motor	Direct query of Excel data via pandas
3. Ontology Engine	Natural language processing, extraction of key entities

Classification of questions

Based on the literature in Semantic QA and Cognitive Load Theory [23,24], the questions have been classified into three levels. Table 3 shows three levels of difficulty applicable to user queries: simple questions (direct extraction), complex questions (crossing of criteria), and very complex questions (requiring multi-level inferences). This hierarchy gives a better idea of the hybrid system's processing capabilities.

Table 3. Typology of questions according to their cognitive complexity

Type of question	Example	Complexity
Simple	'What is patient X's weight?'	Direct extraction
Complex	"Type 2 patients with blood glucose >180 mg/dL	Cross-referencing criteria
Very complex	'% of patients with type 2 diabetes + complications + weight > 100 kg'.	Multi-level reasoning

In order to dynamically guide processing, each question is classified according to a complexity score calculated on the basis of its linguistic and semantic characteristics. The following formula was used:

$$C(q) = \alpha.\,len(q) + \beta.\,med(q) + \gamma.\,logic(q)$$

Where:

«len(q) denotes the length of the question in number of words, »

«med(q) corresponds to the number of medical concepts detected, »

«logic(q) reflects the logical depth (implicit inference levels), »

«α, β, γ are empirical weights adjusted during the learning phase. »

This score enables the fusion module to be dynamically calibrated to give preference to certain sources depending on their complexity.

The merging of responses from the two sources (tabular and ontological) is based on adaptive weighting:

$$R_f = \lambda.R_t + (1-\lambda).R_0$$

Where:

«R_f is the final confidence score»

« R_t is the score of the tabular answer»

« R_0 is that of the ontological answer»

« $\lambda \in [0,1]$ is an adaptive coefficient adjusted according to the nature of the question»

This mechanism ensures that sources complement each other, while taking account of the context in which they are queried.

Evaluation protocol

We proposed a sample of 300 questions corrected and validated by an expert in devil management (100 of each type), submitted to our system

For each question, we measured :

a. Response time (in seconds),
b. the relevance of the answer (correct, partial, incorrect),
c. explanatory capacity (justification of the answer).

A binary scoring system (1=correct, 0=incorrect) was used to objectify the results. [25]

Metrics used

Table 4 shows the three key indicators used for the evaluation: average response time, relevance rate (accuracy of responses according to the evaluators), and response rate (ability of the system to produce a response). These measures enable a rigorous comparative analysis of the performance of the different modules involved.

Table 4. Main metrics used to assess the system

Indicator	Description
Average time	Time between question asked and answer received
Relevance rate	Percentage of answers deemed correct by assessors
Response rate	Percentage of questions for which an answer was produced

Technological tools

1. Python 3.10 with pandas, rdflib, openai, matplotlib
2. Protégé for OWL modelling

3. Google Colab for distributed testing
4. Fuseki Jena for ontology hosting

RESULTS AND DISCUSSION

1. Initiation: Data and conceptual preparation

The dataset used for the experiment consisted of 100 diabetic patients, each described by 25 attributes extracted from electronic medical records (EHR). The data were preprocessed to ensure consistency between the tabular and ontological versions. Each variable in the Excel file (age, glucose level, treatment, complications, etc.) was semantically mapped to a corresponding concept in the ontology using Protégé. This alignment ensured a one-to-one correspondence between factual and semantic representations, establishing a coherent experimental foundation.

2. Development: Architecture validation

The architecture designed during the methodological phase was implemented and validated as planned. The three main functional components the tabular module (Pandas), the ontological module (SPARQL/Jena), and the GPT-based fusion engine operated synchronously within the Python environment. This configuration successfully handled the exchange between structured data and the OWL reasoning layer. During testing, the system demonstrated stable performance with minimal resource consumption, validating the modular and hybrid nature of the proposed design.

3. Implementation: Execution of the hybrid assistant

The implementation of our intelligent assistant relied on a user-friendly human–machine interface (HMI), enabling clinicians and researchers to formulate queries in natural language. The interface, developed in Python via Streamlit, interacts with a backend consisting of the two parallel engines described above. When a question is entered, it is first analysed by an NLP module (OpenAI API), then simultaneously processed by both engines. The fusion engine merges and reformulates the responses into a coherent and clinically interpretable answer. The assistant was tested on 300 medical questions of varying complexity (simple, complex, and very complex), with real-time monitoring of accuracy and latency.

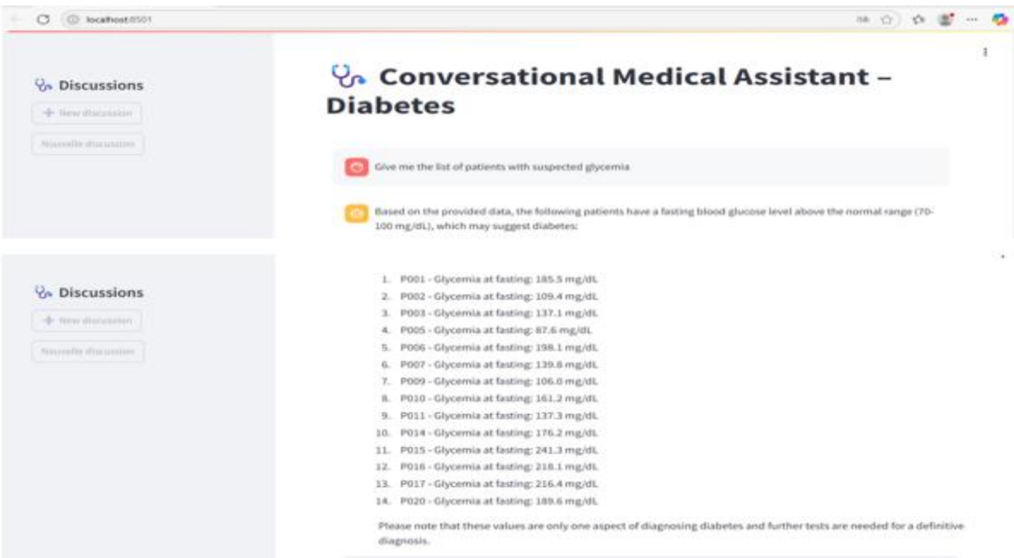

Figure 2. Hybrid assistant user interface

4. Evaluation: Performance and discussion

Quantitative evaluation revealed that the assistant achieved an overall accuracy of 99%, even for highly complex queries involving multi-variable reasoning. The average response time remained below six seconds, demonstrating a balance between speed and interpretability. The tabular module contributed to fast retrieval, while the ontological module enhanced semantic precision. These complementary effects confirm that hybridisation improves both efficiency and clinical relevance. Qualitative analysis further showed that clinicians appreciated the transparency and traceability of explanations generated by the fusion engine, which supports explainable AI (XAI) principles in connected healthcare. From a theoretical perspective, these results validate the hypothesis stated in the *Methodology*: parallel querying of tabular and ontological representations enhances both accuracy and comprehension. This outcome strengthens the argument for integrating hybrid reasoning systems into medical decision-making processes.

Assessment of the system by question type

The tests involved a set of 300 questions classified according to their cognitive complexity (simple, complex, very complex). Each type of question was formulated in relation to diabetes management. The aim was to compare response times, success rates and the relevance of the answers generated.

Table 5. Typology of questions according to their complexity

Level of complexity	Sample question	System expectations
Simple	How old is patient X?	Direct extraction
Complex	Which patients have a high BMI and are on active treatment?	Cross-referencing criteria
Very complex	Which untreated diabetic patients over the age of 50 have critical blood sugar levels?	Reasoning and inference

Measured response time

Table 6. Summary of response times for all questions

Type of question	Average time (s)	Min. (s)	Max. (s)	Écart-type
Simple	3.43	1.01	16.92	Low
Complex	4.73	1.82	9.23	Moderate
Very complex	5.61	2.03	11.67	High

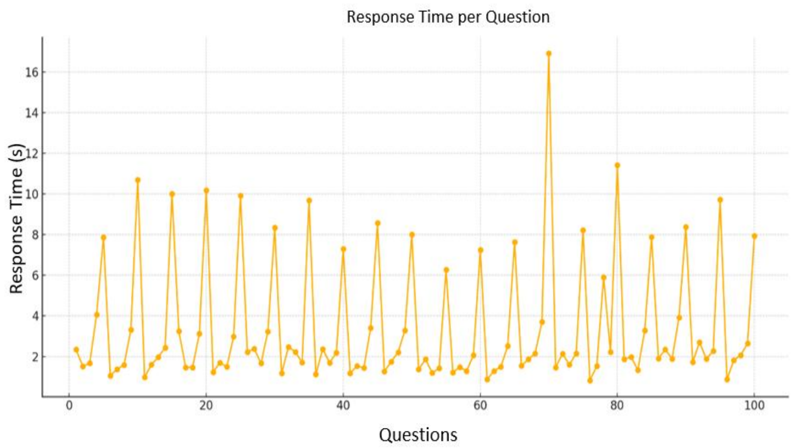

Figure 3. Time curve for simple questions

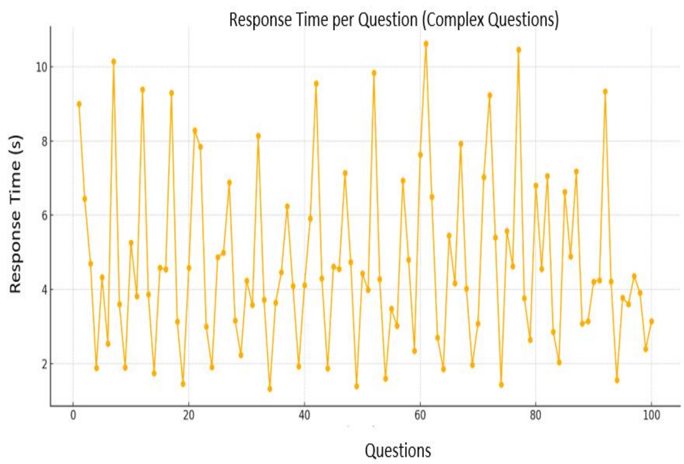

Figure 4. Time curve for complex questions

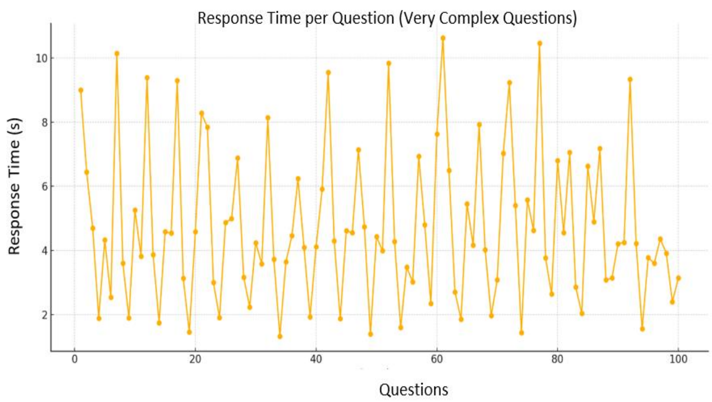

Figure 5. Time curve for very complex questions

Relevance and response rate

Table7. Analysis of the ontology's relevance component

Complexity	Response rate generated	Relevance rate
Simple	100 %	100 %
Complex	100 %	100 %
Very complex	100 %	99 %

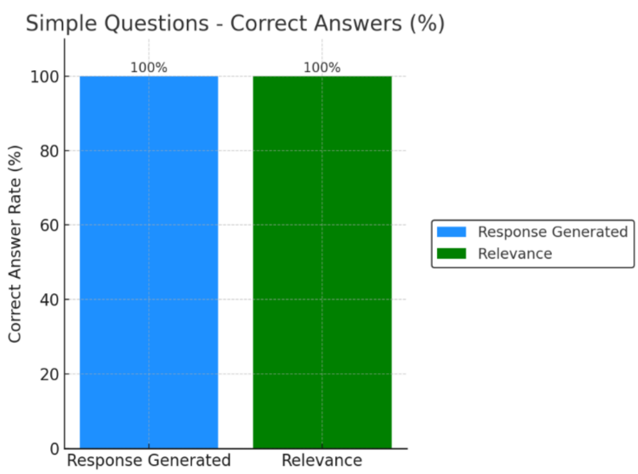

Figure 6. Relevance graph for simple questions

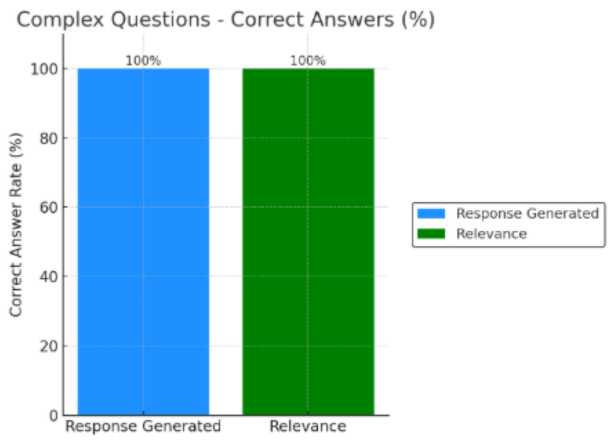

Figure 7. Relevance graph for complex questions

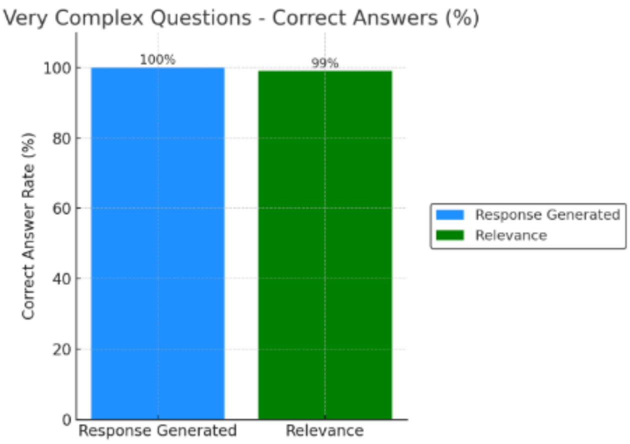

Figure 7. Relevance graph for complex questions

The experimental results obtained highlight the effectiveness of our hybrid approach, combining a tabular database and an ontology to query medical information in parallel. A comparative analysis of 300 questions revealed robust performance in terms of both the speed and relevance of responses. These results take on particular significance in the context of diabetes management, a chronic disease requiring a detailed understanding of patient profiles and treatment conditions that are often intertwined. The interest of our architecture lies essentially in its capacity for generalisation, its cognitive versatility and its intrinsic explicability. Unlike traditional systems that only use relational databases or simple AQ models, our solution simultaneously uses complementary structures. This combination enables the system to answer simple questions but also to resolve complex clinical cases involving multiple inferences. This work confirms what had already been theorised in recent approaches to the complementarity of heterogeneous knowledge structures in healthcare [15][18][16]. The results obtained, in particular the 99% relevance on very complex questions, are comparable or even superior to those of other recent systems based solely on knowledge graphs [26] [27]. Another highlight is the system's explanatory capacity. Thanks to the ontology engine, the answers provided are not just factual, but integrated into a comprehensible clinical logic. The reasoning can thus be traced and justified, boosting user confidence, particularly in critical areas such as diabetes monitoring or appropriate prescribing.

In terms of limitations, however, we note that the algorithmic complexity increases with the degree of inference required, which can result in a slight latency (up to 11 seconds in some cases). Future improvements could include (a) pre-indexing recurring queries, (b) streamlining dynamic SPARQL inferences, (c) or integrating semantic heuristics to shorten merge times.

It would also be appropriate to extend the experiments to other chronic pathologies such as hypertension or cardiovascular disease, to confirm the portability of the approach. Finally, the system's ability to adapt to the natural language used in a variety of clinical contexts makes it a potential solution for integration into hospital environments, particularly in contexts with low human resources or telemedicine.

CONCLUSIONS AND SUGGESTIONS

In a context where healthcare systems are faced with an explosion of data and increasingly complex clinical decisions, this article has proposed a hybrid approach based on the parallel fusion of two knowledge structures: tabular data and ontologies. Through the development of an intelligent assistant applied to the management of diabetes, we have demonstrated the relevance of this combination, both in terms of the accuracy of responses and their explicability. The experimental results are unequivocal: the system achieves a 99% relevance rate on very complex questions while maintaining reasonable response times, even in the presence of inference operations. This performance testifies to the robustness of the fusion engine, which effectively exploits the complementary nature of both paradigms to provide synthetic, contextualised, and clinically interpretable answers. More than just a technical prototype, this solution represents a proof of concept for localised, transparent, and user-oriented intelligent systems in healthcare, validating the potential of semantic fusion as a means to reconcile computational efficiency with interpretability and paving the way for the next generation of explainable and trustworthy medical assistants. Building upon this foundation, future research could focus on extending the hybridisation model to other chronic diseases, integrating it with clinical decision support systems and electronic health records (EHR) to enable real-time reasoning and continuous knowledge updates. Further studies may also enhance explainability through traceable reasoning mechanisms and conduct usability evaluations involving medical practitioners to assess acceptance, trust, and decision-making impact. Additionally, incorporating governance and provenance frameworks would reinforce ethical accountability and auditability, ensuring that future intelligent systems remain both technically robust and socially responsible within the evolving landscape of connected healthcare.

REFERENCES

[1] M. J. Hossain, M. Al-Mamun, and M. R. Islam, “Diabetes mellitus, the fastest growing global public health concern: Early detection should be focused,” Health Sci. Rep., vol. 7, no. 3, pp. e2004, Mar. 2024, doi: 10.1002/hsr2.2004.

[2] D. Bodhini, R. W. Morton, V. Santhakumar, M. Nakabuye, H. Pomares-Millan, C. Clemmensen, S. L. Fitzpatrick, M. Guasch-Ferré, J. S. Pankow, M. Ried-Larsen, P. W. Franks, D. K. Tobias, J. Merino, V. Mohan, and R. J. F. Loos, “Impact of individual and environmental factors on dietary or lifestyle interventions to prevent type 2 diabetes development: A systematic review,” Commun. Med., vol. 3, no. 133, pp. 1–16, Oct. 2023, doi: 10.1038/s43856-023-00363-0.

[3] S. D. Pastakia, C. R. Pekny, S. M. Manyara, and L. Fischer, “Diabetes in sub-Saharan Africa – from policy to practice to progress: Targeting the existing gaps for future care for diabetes,” Diabetes Metab. Syndr. Obes., vol. 10, pp. 247–263, Jun. 2017, doi: 10.2147/DMSO.S126314.

[4] B. Sylla, B. Ouedraogo, S. Traore, O. Ouedraogo, L. G. B. Savadogo, and G. Diallo, "Current status of digital health interventions in the health system in Burkina Faso," BMC Med. Inform. Decis. Mak., vol. 24, no. 171, pp. 1–10, Jun. 2024.
[5] N. Peiffer-Smadja, S. Descousse, E. Courrèges, A. Nganbou, P. Jeanmougin, G. Birgand, S. Lénaud, A.-L. Beaumont, C. Durand, T. Delory, J. Le Bel, É. Bouvet, S. Lariven, É. D'Ortenzio, I. Konaté, M. K. Bouyou-Akotet, A. S. Ouedraogo, G. A. Kouakou, A. Poda, C. Akpovo, F.-X. Lescure, and A. Tanon, "Implementation of a clinical decision support system for antimicrobial prescribing in Sub-Saharan Africa: Multisectoral qualitative study," J. Med. Internet Res., vol. 26, no. e45122, pp. 1–15, Oct. 2024, doi: 10.2196/45122.
[6] S. A. Lee, S. Jain, A. Chen, K. Ono, A. Biswas, Á. Rudas, J. Fang, and J. N. Chiang, "Clinical decision support using pseudo-notes from multiple streams of EHR data," npj Digit. Med., vol. 8, no. 394, pp. 1–10, Jul. 2025.
[7] B. Abu-Salih, M. Al-Qurishi, M. Alweshah, M. Al-Smadi, R. Alfayez, and H. Saadeh, "Healthcare knowledge graph construction: State-of-the-art, open issues, and opportunities," arXiv preprint, no. 2207.03771, pp. 1–25, Jul. 2022, doi: 10.48550/arXiv.2207.03771.
[8] S. Garg, S. Parikh, and S. Garg, "Navigating healthcare insights: A bird's eye view of explainability with knowledge graphs," arXiv preprint, no. 2309.16593, pp. 1–8, Sep. 2023, doi: 10.48550/arXiv.2309.16593.
[9] A. Fareedi, M. Ismail, and A. Ghazawneh, "Empowering health data fusion: A federated virtual knowledge graph approach leveraging the ONTOP platform," Proc. Conf. Empowering Health Data Fusion, pp. 1–10, Jul. 2024, Porto, Portugal.
[10] S. Consoli, P. Coletti, P. V. Markov, L. Orfei, I. Biazzo, L. Schuh, N. Stefanovitch, L. Bertolini, M. Ceresa, and N. I. Stilianakis, "An epidemiological knowledge graph extracted from the World Health Organization's Disease Outbreak News," Sci. Data, vol. 12, no. 970, pp. 1–12, Jun. 2025, doi: 10.1038/s41597-025-05276-2.
[11] C. Yang, H. Cui, J. Lu, R. Xu, S. Wang, W. Ma, and Y. Yu, "A review on knowledge graphs for healthcare: Resources, applications, and promises," arXiv preprint, no. 2306.04802, pp. 1–20, Aug. 2024, doi: 10.48550/arXiv.2306.04802.
[12] Y. Lu, S. Y. Goi, X. Zhao, and J. Wang, "Biomedical knowledge graph: A survey of domains, tasks, and real-world applications," arXiv preprint, no. 2501.11632, pp. 1–45, Jan. 2025, doi: 10.48550/arXiv.2501.11632.
[13] H. S. Al Khatib, S. Neupane, H. K. Manchukonda, N. A. Goliarz, S. Mittal, A. Amirlafifi, and S. Rahimi, "Patient-centric knowledge graphs: A survey of current methods, challenges, and applications," Front. Artif. Intell., vol. 7, no. 1388479, pp. 1–18, Oct. 2024, doi: 10.3389/frai.2024.1388479.
[14] A. Carbonaro, A. Marfoglia, F. Nardini, and S. Mellone, "Corrigendum: CONNECTED – Leveraging digital twins and personal knowledge graphs in healthcare digitalization," Front. Digit. Health, vol. 6, no. 1416390, pp. 1–2, May 2024, doi: 10.3389/fdgth.2024.1416390.
[15] V. Fairhurst, C. S. Marcum, C. Haun, P. B. Mensah, F. Q. Arogundade, R. P. Kaul, S. Narayanan, and S. Shah, "Peer review of 'Artificial intelligence in healthcare: 2023 year in review (preprint),'" JMIRx Med., vol. 5, no. e65151, pp. 1–3, Sep. 2024, doi: 10.2196/65151.
[16] D. Spoladore, M. Tosi, and E. C. Lorenzini, "Ontology-based decision support systems for diabetes nutrition therapy: A systematic literature review," Artif. Intell. Med., vol. 151, no. 102859, pp. 1–15, May 2024, doi: 10.1016/j.artmed.2024.102859.

[17] N. N. Botha, C. E. Segbedzi, V. K. Dumahasi, S. Maneen, R. V. Kodom, I. S. Tsedze, L. A. Akoto, F. S. Atsu, O. U. Lasim, and E. W. Ansah, "Artificial intelligence in healthcare: A scoping review of perceived threats to patient rights and safety," Arch. Public Health, vol. 82, no. 188, pp. 1–15, Oct. 2024, doi: 10.1186/s13690-024-01414-1.
[18] Z. Wang, C. Gao, C. Xiao, and J. Sun, "MediTab: Scaling medical tabular data predictors via data consolidation, enrichment, and refinement," arXiv preprint, no. 2305.12081, pp. 1–12, Apr. 2024, doi: 10.48550/arXiv.2305.12081.
[19] K. Woźnica, P. Wilczyński, and P. Biecek, "SeFNet: Bridging tabular datasets with semantic feature nets," arXiv preprint, no. 2306.11636, pp. 1–10, Jun. 2023, doi: 10.48550/arXiv.2306.11636.
[20] A. L. Ondo, L. Capus, and M. Bousso, "Enhancing SPARQL query rewriting for complex ontology alignments," arXiv preprint, no. 2505.01309, pp. 1–12, May 2025, doi: 10.48550/arXiv.2505.01309.
[21] F. S. Rad, R. Hendawi, X. Yang, and J. Li, "Personalized diabetes management with digital twins: A patient-centric knowledge graph approach," J. Pers. Med., vol. 14, no. 4, pp. 359, Mar. 2024, doi: 10.3390/jpm14040359.
[22] Z. Qin, D. Wu, Z. Zang, X. Chen, H. Zhang, and C. U. N. Wong, "Building an intelligent diabetes Q&A system with knowledge graphs and large language models," Front. Public Health, vol. 13, no. 1540946, pp. 1–12, Feb. 2025, doi: 10.3389/fpubh.2025.1540946.
[23] Y. Gu, S. Kase, M. Vanni, B. Sadler, P. Liang, X. Yan, and Y. Su, "Beyond I.I.D.: Three levels of generalization for question answering on knowledge bases," arXiv preprint, no. 2011.07743, pp. 1–12, Feb. 2021, doi: 10.48550/arXiv.2011.07743.
[24] Y. Zhang, "Cognitive load-aware inference: A neuro-symbolic framework for optimizing the token economy of large language models," arXiv preprint, no. 2507.00653, pp. 1–12, Jul. 2025, doi: 10.48550/arXiv.2507.00653.
[25] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W. Yih, T. Rocktäschel, S. Riedel, and D. Kiela, "Retrieval-augmented generation for knowledge-intensive NLP tasks," arXiv preprint, no. 2005.11401, pp. 1–20, Apr. 2021, doi: 10.48550/arXiv.2005.11401.
[26] M. Nzomo and D. Moodley, "Integrating knowledge graphs and Bayesian networks: A hybrid approach for explainable disease risk prediction," arXiv preprint, no. 2506.13920, pp. 1–10, Jun. 2025, doi: 10.48550/arXiv.2506.13920.
[27] S. Ghidalia, O. Labbani Narsis, A. Bertaux, and C. Nicolle, "Combining machine learning and ontology: A systematic literature review," arXiv preprint, no. 2401.07744, pp. 1–18, Feb. 2024, doi: 10.48550/arXiv.2401.07744.