Analecta Technica Szegedinensia
Vol. 19, No. 2 ISSN 2064-7964 2025

SEMANTIC FUSION FOR CONNECTED MEDICINE: AN EXPERIMENTAL COMPARISON OF TABULAR AND ONTOLOGICAL STRUCTURES IN THE EFFICIENT MANAGEMENT OF DIABETES

Muhala Luhepa Blaise[1],Munduku Munduku Deo[1], Bukasa Kakamba John[2], Oshasha Oshasha Fiston [1], Mabela Makengo Rostin[1], Muluba Celestin[1], Munene Asidi Djonive[3]*

[1]Department of Mathematics, Statistics and Computer Science, University of Kinshasa, Kinshasa, DR Congo
[2]Endocrinology, Metabolism, and Nuclear Medicine, Kinshasa University Clinics, Kinshasa
[3]Department of Exploration and Production, Faculty of Oil, Gas and New Energies, University of Kinshasa, DR Congo.

e-mail: blaise.muhala@unikin.ac.cd

Received: 21st June	*Accepted: 26th October*

ABSTRACT

The explosion in digital medical data makes it crucial to design intelligent assistants capable of interrogating this information reliably, quickly and contextually. In this article, we propose a novel comparative approach between two forms of knowledge representation: traditional tabular data and semantic ontologies. Based on the same clinical dataset concerning diabetic patients, we have implemented a dual structuring: a tabular version and an RDF ontology modelled with Protégé. An intelligent assistant, interfaced with the GPT-4 API, was designed to query both formats. The originality of our contribution lies in the experimental parallelisation of these two data models, through a standardised series of 300 questions, classified according to three levels of increasing complexity. This methodology enables us to objectively assess the robustness, responsiveness and inference capacity of each approach. The results are unequivocal: the ontology systematically outperforms the tabular format, with exact response rates ranging from 97% to 100%, compared with 34% to 81% for the tabular format. In addition, the ontological approach shows better tolerance of ambiguous queries and stability in semantic interpretation. Over and above performance, this study highlights the potential of knowledge graphs as an architectural foundation for future medical decision support systems. It also paves the way for hybrid systems that combine the accessibility of tables with the semantic power of ontologies - a perspective that has so far been little explored in the context of connected healthcare.

Keywords: Medical ontology, Intelligent assistant, Knowledge representation, Connected medicine, XLSX/OWL comparison

1. INTRODUCTION

Diabetes is a heterogeneous metabolic disease characterised by chronic hyperglycaemia due to impaired insulin secretion, impaired insulin action, or both. It is classified as type 1 diabetes, type 2 diabetes, gestational diabetes and other types of diabetes. Diabetes has become a major public health problem because of its prevalence, its complications and the cost of treating them. The incidence of diabetes has risen sharply over the last 30 years, from 7% in 1990 to 14% in 2022. However, in 2019, the International Diabetes Federation projected that 783 million diabetics would be affected by 2045, and here we are, having already exceeded this estimate, with 830 million diabetics by 2022.[1],[2].This increase is particularly worrying in low- and middle-income countries, where healthcare systems are often ill-suited to providing continuous, personalised care. The situation in Africa is alarming. Some 54 million people in Africa are currently living with diabetes, and this figure could double in the next two decades [3]. It has been shown that it is the poor countries of Africa and Asia that will be the source of a large number of diabetics in the years to come, as a result of stress and an unbalanced diet. More than half of cases are neither diagnosed nor treated, exposing patients to severe complications, particularly cardiovascular, neurological and ophthalmological [4]. There are many obstacles to this: lack of qualified staff, absence of standard protocols, poor integration of medical data and high cost of treatment [5].

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

In the Democratic Republic of Congo, according to the International Diabetes Federation (IDF), in 2024, the prevalence of diabetes among adults (aged 20-79) in the DRC is estimated at 7.7%, which represents around 2.86 million cases [22], yet the country has only 5 endocrinologists for a population of almost 120 million, patients are not sufficiently educated, there are no structures to encourage physical activity to combat sedentary lifestyles, there are no technical facilities up to scratch, and there is no universal healthcare cover. Patients are crying out for help from a system that cannot guarantee them a better tomorrow. Faced with these challenges, technology could play an important role in improving care. In particular, intelligent assistants capable of interrogating medical databases appear to be promising solutions for improving monitoring, diagnosis and clinical decision-making [6], [7]. However, the effectiveness of these assistants depends heavily on the structure of the data they use. Tabular data (Excel, CVS, JSON, etc.) are still widely used because they are easy to access, but they have significant limitations in terms of relational logic and semantic interpretation [8]. For example, in an Excel file, it is difficult to represent implicit relationships such as ‘a patient is at risk in the absence of treatment and the presence of a complication’ [9]. This is where medical ontologies, particularly those described in OWL (Web Ontology Language), come into their own. They enable knowledge to be structured in the form of triplets (subject-predicate-object) while capturing hierarchical, causal and contextual relationships [10], [11]. This formalism facilitates automatic reasoning using inference engines or SPARQL queries [12], [13]. The integration of ontologies with large language models (LLMs), such as GPT-4, offers even more powerful prospects. Thanks to their ability to understand natural language, LLMs can be coupled with knowledge graphs to provide contextualised and justifiable answers [14], [15]. Recent work has shown that fine-tuning these models with specialised medical corpora considerably improves the accuracy of the answers provided by the assistants [16], [17]. For example, Doumanas et al. have shown that GPT-4, when trained on ontological structures, can automatically generate coherent medical vocabularies [18]. The DRAGON-AI project has proposed a method for dynamically generating ontologies from clinical texts, facilitating the construction of reusable knowledge graphs [19]. Initiatives such as the Swiss Personalized Health Network (SPHN) and the Care and Registry Semantic Model (CARE-SM) have integrated these approaches into their information systems to improve the interoperability of healthcare data [20], [21]. Despite these advances, few studies have experimentally compared the performance of intelligent assistants using tabular versus ontological structures. This is precisely the aim of this study. Using a dataset of diabetic patients, we designed an intelligent assistant capable of querying both tabular and ontological data formats in order to observe the differences in performance as a function of the complexity of the questions asked. For our experiments we used a dataset in xlsx format and an ontology in OWL format.Three series of 100 questions (simple, complex and very complex). The 300 medical questions were designed on the basis of the diabetes management standards established by the Congolese health authorities [23] and enriched by recent scientific recommendations on diabetes in Africa [24] [25]. For each series, we evaluated the relevance of the responses, the error rate and the average treatment time. The final objective is to demonstrate that ontologies, because of their logical and semantic structuring, are better suited to the development of intelligent medical assistance systems, particularly in sensitive contexts such as diabetes management.

1.1. State Of The Art

Several recent research projects have provided food for thought and guided the development of our approach, particularly in the field of semantic representation applied to diabetes monitoring. Zhou et al looked at the integration of ontologies into clinical decision support systems specifically designed for diabetes. Their study showed that formalised knowledge bases not only improved the accuracy of medical recommendations, but also enhanced the traceability of diagnostic decisions in complex situations [26]. In a complementary study, Flory et al (2025) compared the responses of the GPT-4 model with those of a group of 31 endocrinologists concerning the initial choice of treatment for diabetic patients. The results highlighted the model's ability to propose decisions comparable to those of clinicians, particularly in

Analecta Technica Szegedinensia
Vol. 19, No. 2 ISSN 2064-7964 2025

contexts where the data are ambiguous or partial [27].Doumanas et al addressed a technical problem: how to use a large language model to automatically generate structured ontologies. They designed a pipeline exploiting GPT-4 to construct coherent graphs from textual descriptions, thus opening the way to hybrid forms of modelling combining linguistic intelligence and ontological rigour [18]. In the same spirit of semantic automation, the DRAGON-AI project, led by Toro et al, focused on the dynamic generation of ontologies from unstructured medical narratives. Using neural models to interpret these texts, they have developed a system capable of transforming medical records into usable knowledge graphs [28]. From an application point of view, Seneviratne et al. evaluated a device for simulating semantic queries in the field of clinical decision-making, in particular in the context of diabetic comorbidities. Their RDF platform was used to test intelligent agents on their ability to infer relevant responses from linked data [29]. Furthermore, the issue of interoperability in low-resource healthcare environments was at the heart of the study conducted by Palojoki et al. The authors demonstrate that the transition from traditional tabular representations to semantic structures not only improves the quality of medical reasoning, but also facilitates the integration of heterogeneous data from different systems [30]. In West Africa, Nacanabo et al. conducted an in-depth investigation into digital literacy and the challenges associated with the adoption of diabetes monitoring tools in low-infrastructure settings. Their study shows a persistent digital divide which hinders the appropriation of decision-support technologies [31]. Finally, Ouedraogo's team proposed a summary of the obstacles encountered in the management of type 2 diabetes in primary care in West Africa. Their review highlights the central role that intelligent semantic assistants could play in coordinating care pathways, detecting at-risk cases and recommending targeted actions [32]. Although each of these studies sheds light on a specific aspect of semantic integration in healthcare, none to date has experimentally compared a conventional tabular dataset with its ontological equivalent, which can be queried using the same intelligent assistant. It is precisely this gap that we are seeking to fill with our approach.

2. METHODOLOGY

2.1. Comparative experimental approach

The methodology adopted in this article is based on a comparative experimental approach. Its objective is to evaluate the impact of the ontological and tabular knowledge representation format on the performance of an intelligent assistant responsible for answering medical questions from the same diabetes dataset. [33] [34] Its architecture is shown in figure 1.

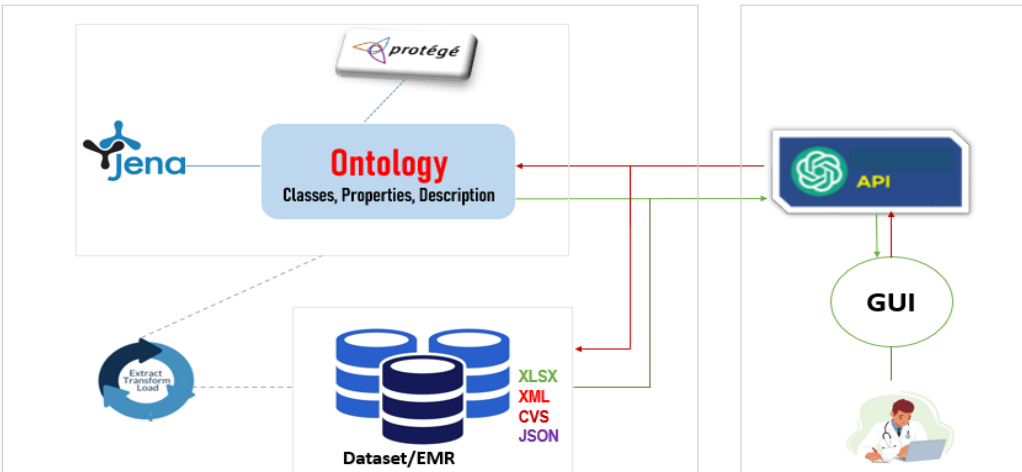

Figure 1: Architecture of our intelligent assistant system

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

Analecta Technica Szegedinensia
Vol. 19, No. 2 ISSN 2064-7964 2025

2.2. General design of the study

This study is based on a comparative methodology aimed at empirically evaluating the relevance and performance of two approaches to representing medical knowledge in the context of connected medicine: a tabular approach (based on a structured dataset) and an ontological approach. The main objective is to determine whether the semantic structuring of data via an ontology improves the quality of responses produced by an intelligent assistant, compared with non-semantically enriched data [26], [18].
The experimental design is based on a mirrored protocol, where the same clinical dataset was queried using two modalities:
(1) direct interrogation of the dataset using filters and conditional formulas in Python (Pandas),
(2) semantic interrogation of the same tabular dataset transformed into an ontology using SPARQL queries via Jena Fuseki [35].Table 1 below provides a comparative summary of the two representation and interrogation paradigms used in this study.

Table 1: Summary comparison of the two experimental methods in terms of format, knowledge structure, inference capacity and interrogation tools.

Criteria	Tabular approach	Ontological approach
Data format	Structured spreadsheet (.xlsx)	RDF/OWL Ontology
Query language	Conditional formulas(Pandas)	SPARQL
Knowledge structure	Linear, flat	Hierarchical, relational
Capacity for inference	Limited	High (via reasoning)
Main tool	Python Pandas	Jena Fuseki + RDFLib
Interoperability	Low	High

To ensure rigorous and consistent evaluation, an intelligent conversational assistant has been developed to act as a questioning interface. This is coupled to the GPT-4 API, which is responsible for automatically reformulating natural language questions into formal queries adapted to each data structure (formulas or SPARQL). Interaction is via a medical user-oriented human-machine interface (HMI) [36], [37].
The design of this study is based on four fundamental methodological pillars:
- **Standardisation of input data:** a dataset in xlxs format, containing data from 100 diabetic patients, was used as a starting point. This file was cleaned, structured and then exported in RDF/OWL to ensure perfect correspondence between the two formats.
- **Symmetry of questioning:** a corpus of 300 questions covering simple to complex medical cases was developed. These questions were formulated independently of the data format, then injected into the wizard for double execution (XLSX vs OWL).
- **Automated processing:** all interactions, from reformulation to response, were automated to avoid any human bias. The system generates queries, executes searches, times response times and records results.
- **Traceability and comparative analysis:** for each question, the results of the two approaches were recorded (raw response, validity, response time) and then compared quantitatively and qualitatively using objective indicators. This system was designed to answer a central question: does an ontology really improve an assistant's ability to provide relevant, rapid and interpretable medical responses based on local data?

2.3. Data sources

The dataset used in this study is based on a representative clinical database constructed from the information of 100 diabetic patients collected at the Centre Hospitalier HN, located in the commune of Mont-Ngafula in Kinshasa. This source file, prepared in Microsoft Excel (.xlsx), reflects the types of data commonly observed in electronic medical records (EMRs), incorporating biometric, therapeutic and

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

Analecta Technica Szegedinensia
Vol. 19, No. 2 ISSN 2064-7964 2025

clinical variables essential to the analysis. Each record (row) represents a unique patient, and each column corresponds to a key medical property. Attributes included include:
- Patient identifier ;- Sex, age ;- Type of diabetes ;- Blood glucose (fasting, postprandial) ;- Body mass index (BMI);- Current treatment (insulin, oral antidiabetics, diet alone, etc.);- Presence of complications (peripheral neuropathy, autonomic neuropathy, retinopathy, nephropathy, ischaemic heart disease, ischaemic stroke, obliterative arterial disease of the lower limbs, fri);- Most recent follow-up date.

The data was selected to cover both simple clinical cases (direct question on one value) and more complex cases involving several cross-referenced factors. To enable a fair comparison with a semantic environment, this xls format file was manually transformed into an OWL ontology. Modelling was carried out in Protégé 5.6, using ontology engineering principles: structuring into classes (Patient, TypeDiabete, imbalance or revelation factor, complication, treatment, etc.), relationships (aTreatment, presentComplication, aGlycemia, etc.) and instances. Each xlsx record was converted into RDF triples that could be used by reasoning engines such as Jena. Figure 2 illustrates this transformation between a tabular record in xlsx (Excel) format and its semantic equivalent represented in RDF/OWL, as used in our experimental model.

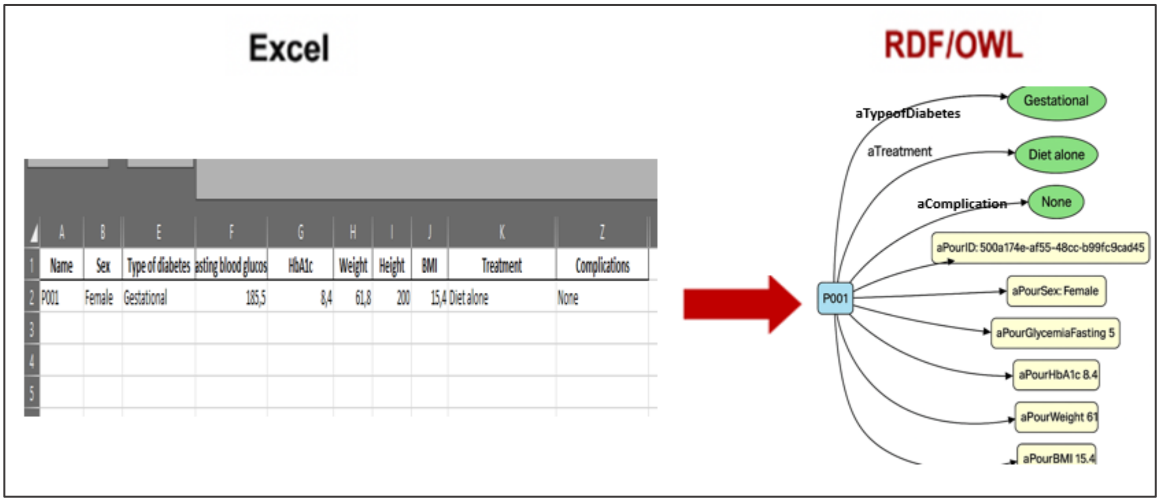

Figure 2: Illustration of the conversion of an Excel record to an RDF/OWL model.

This dual structuring ensures that the two formats contain exactly the same information, but encoded according to two radically different paradigms. In this way, any difference in the performance of the intelligent assistant will be directly attributable to the knowledge representation model, and not to the data itself. The use of ontologies in clinical decision support systems (CDSS) has been shown to improve the automation and transparency of the reasoning process, facilitating the generation of interpretable and accurate treatment recommendations [26]. Studies have explored the potential of large language models (LLMs) to automate the generation of OWL ontologies from natural language descriptions, introducing new elicitation techniques for automated ontology development [36]. Furthermore, ontology engineering plays a crucial role in structured knowledge modelling and management, with research evaluating the performance of language models such as GPT-4 and Mistral 7B in efficiently automating ontology engineering tasks [18]. Finally, the interoperability of electronic health records is essential to improve care coordination and patient outcomes, with conceptual frameworks proposed to address the associated technical and semantic challenges [35].

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

2.4. Implementation of the intelligent assistant

To ensure consistent and automated interaction with the two representation models (XLSX and OWL), an intelligent assistant has been designed. This assistant is based on a modular architecture built around three main components: (1) a question entry interface, (2) a linguistic interpretation engine based on the OpenAI GPT-4 API, and (3) two independent query modules: one for the XLSX format dataset, the other for the RDF/OWL graphs.

a) Interpreting natural language queries

The user interacts with the system via a natural language interface, simulating a doctor-assistant interface. The question entered is transmitted to GPT-4 via its API, with an explicit instruction to produce either a Python/Pandas formula for tabular processing or a SPARQL query for ontological processing, depending on the context selected. Studies have shown that GPT-4 can improve the accuracy of clinical decisions, in particular by assisting doctors in specialised fields such as nephrology [38]. In addition, GPT-4 has demonstrated its ability to generate precise SPARQL queries from natural language questions, thus facilitating the interrogation of medical knowledge graphs [39]. Figure 3 below illustrates an interaction with the interface developed. The assistant, powered by GPT-4, reformulates the user query in natural language and returns a medically contextualised response, automatically extracted from the data file.

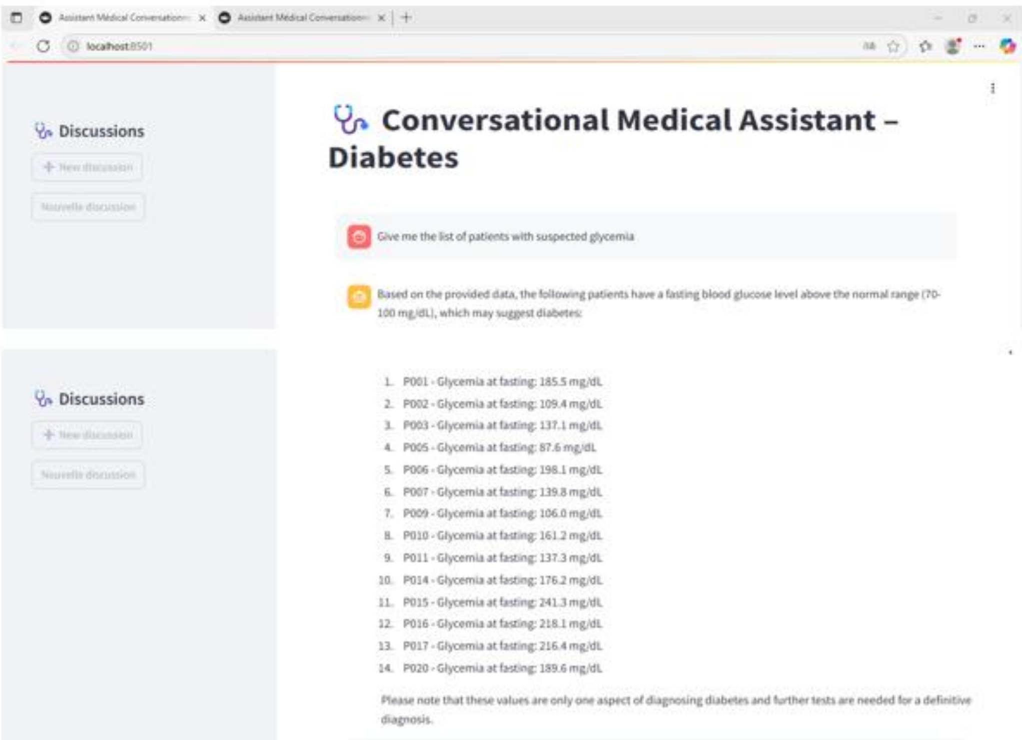

Figure 3: Conversational medical assistant interface. The user enters a question in natural language ('Give me a list of patients with suspect blood sugar'); the system identifies the medical criterion (blood sugar > 126 mg/dL) and returns a structured response extracted from the data file.

b) Excel processing via Python

The Excel module is based on the Pandas library, which loads the .xlsx database, applies filters, aggregates data or returns targeted values. The process is sequential and deterministic.
The use of Pandas for healthcare data analysis is well established, offering powerful tools for cleaning, manipulating and analysing clinical datasets [40].

c) Ontological processing via SPARQL

The ontology module is based on two tools:
- RDFLib for loading and manipulating OWL files locally, - SPARQLWrapper for querying a remote triplestore (via Jena Fuseki). The SPARQL queries generated by GPT-4 are executed dynamically, and the results are formatted for display in a way that is readable by the user. Research has evaluated the ability of large language models to generate valid SPARQL queries, highlighting their potential in querying complex knowledge graphs [38].

d) Timing and recording of results

The following information is automatically recorded for each question asked:- the format used (Excel or OWL),- the query generated by GPT-4,- the response returned,- the processing time (measured by time.time()),- relevance assessment (correct, incorrect, partially correct).
All this data is fed into a database of results used in section 4 for comparative statistical analysis.
Figure 4 below summarises the entire process of handling a query by the intelligent assistant, from receiving the question in natural language to providing the answer via the user interface, via the choice of engine, interpretation by GPT-4, execution and reformulation of the answer.

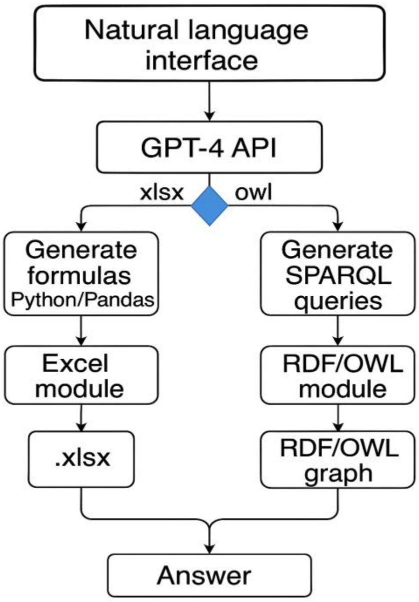

Figure 4: Functional flowchart of the conversational medical assistant. This diagram describes the entire process of handling a medical question, from input to response, depending on whether the database queried is tabular (Excel) or semantic (OWL).

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

2.5. Construction of the test instrument

In order to rigorously evaluate the performance of the intelligent assistant, a standardised test instrument was developed, comprising a set of 300 questions. These questions were classified according to three levels of complexity, defined as follows:
- **Simple questions:** These questions involve a single entity or criterion, without requiring complex reasoning. They can be resolved by direct information extraction. For example: "What is the weight of patient P014?
- **Complex questions:** These require a combination of several criteria or cross-filtering. For example: "Which patients have blood sugar levels > 180 and are not taking insulin?
- **Very complex questions:** These questions require aggregation operations, statistical calculations or advanced logical inferences. For example: "What percentage of type 2 patients on insulin treatment have a complication?

This classification is inspired by recent work on querying complex knowledge bases, which distinguishes between simple and complex questions based on the number of entities involved, the multiple relationships and the logical operations required to answer them [38].

a) Breakdown of questions

300 questions from the test sample were developed manually from the clinical dataset. They were divided into three categories according to their level of complexity, in line with a typology used in the literature on question-answer systems and semantic mining. Table 2 below shows this breakdown

Table 2: Breakdown of the 300 questions used to assess the assistant according to three levels of complexity.

Level	Number of questions	Description
Simple	100	Direct questions on a single property
Complex	100	Queries with multiple conditions (e.g. AND, OR, NOT)
Very complex	100	Queries requiring group, calculation or inference operations

b) Development of the validation framework

We have created a master reference file, containing for each question:

-The question ID; -The level of complexity; -The expected answer;-The target parameters (column or property targeted);-An automatic evaluation field (0 = false, 1 = correct).

This repository was used as a basis for comparing the answers given by the wizard with the actual results extracted manually.

c) Performance metrics collected (formal definitions)

The evaluation of intelligent assistant performance is based on standardised metrics, widely used in question and answer (QA) systems and validated by recent studies.

2.6. Relevance of the response (Accuracy)

Relevance measures whether the response generated by the system corresponds to the expected reference response. It is defined by :

$$Pertinence = \frac{N_{correct}}{N_{total}}$$

- $N_{correct}$ – Number of correct answers
- N_{total} – Total number of questions evaluated

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

This metric is commonly used to evaluate QA systems, as highlighted in a study on the evaluation of ChatGPT as a question-answer system [38]. It was with this in mind that we implemented our heuristic, which consists of weightings of 1 for correct answers and 0 for incorrect answers.

2.7. Average Response Time

The average processing time per question is given by :

$$\bar{T} = \frac{1}{N} \sum_{i=1}^{N} T_i$$

T_i −Response time for question i (in seconds)
N −Total number of questions answered

This metric is essential for assessing the responsiveness of the system, as discussed in recent guides on performance metrics for AI models [41].

2.8. Success rate by level of complexity

For each category of question (simple, complex, very complex), the success rate is :

$$Rates_{success}^{(k)} = \frac{N_{correct}^{(k)}}{N_{total}^{(k)}} \times 100$$

k ∈ {simple,complex,very complex}

This approach enables a detailed analysis of performance according to the complexity of the questions, based on established practices in the evaluation of QA systems [42].

2.9. xlxs vs owl concordance rate

This rate measures the proportion of questions for which the two systems give the same answer:

$$Concordance = \frac{N_{same\ answers}}{N_{total}} \times 100$$

d) Automation of the experimental protocol

An automated script in Python ensures :
- Reads the 300 questions per series of 100 from a .txt file,- Automatic transmission to the API (GPT-4) with specific instructions,- Reformulation into a Pandas formula or SPARQL query,- Execution of the query,- Record the result, response time and format,- Comparison with the repository,- Saving to a log file.
The experimental process set up to evaluate the intelligent assistant is based on a complete automation loop, with no human intervention. The experimental pipeline shown in Figure 5 below illustrates the sequential stages of execution, from reading the questions to collecting and recording the results. This protocol guarantees the reproducibility, rigour and objectivity of the evaluation.

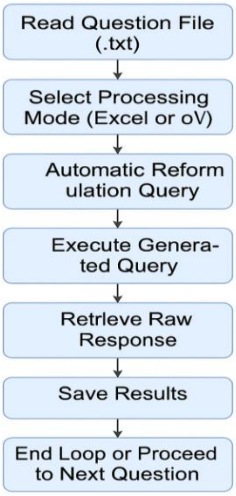

Figure 5: Automation pipeline for the experimental protocol. Each question is extracted from a text file, reformulated automatically by GPT-4 according to the type of data (Excel or OWL), run on the appropriate engine (Pandas or SPARQL) and then evaluated in real time. The results are saved for comparative analysis.

2.10.Assessment criteria

In order to objectively compare the performance of the intelligent assistant according to the two interrogation methods (Excel vs. OWL), a set of evaluation criteria was defined. These criteria aim to measure not only the accuracy of the responses, but also their speed, their robustness in the face of complexity, and their ability to produce a response that can be used by a healthcare professional.

a) Accuracy of answers

Each response generated is compared with a validated reference response. Two levels are used:
- Correct response: exact match with the reference ;- Incorrect response: content or target error

The evaluation is binary (1/0) depending on statistical requirements.

b) Complexity tolerance

This criterion measures the robustness of the system according to the level of difficulty of the question (simple, complex, very complex). A robust system should maintain a high success rate, even on the most demanding cases.

c) Average response time

Processing time is measured for each question from the moment the API call is triggered until the structured response is received. This time is then averaged by level of complexity and by basic format. It reflects the system's operational efficiency.

d) Concordance rate between Excel and OWL

For each question, the wizard queries the xlsx format dataset and the owl ontology database separately. The results are compared to assess their consistency. In particular, this helps to identify cases where :
- Both systems give the same answer,- Only one system responds correctly,- Both fail.

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

This concordance rate is crucial for measuring the potential ontological advantage in complex or ambiguous cases.

e) Linguistic quality of the response

Even if the response is correct, it must be intelligible to a non-technical user (doctor, nurse, etc.). This criterion measures the system's ability to render a clear and legible response, with a natural, complete and unambiguous formulation. Table 3 below summarises these criteria and the methods used to measure them.

Table 3: Summary of the criteria used to benchmark the intelligent assistant. These indicators measure the accuracy of responses, their speed, their adaptability to complexity, their inter-format consistency and their final readability.

Criteria	Description	Type of measurement
Accuracy of answer	Check that the answer matches the reference answer	Binary or weighted score (1/0/0.5)
Tolerance of complexity	Evaluates the ability to respond according to level (simple, complex, very complex)	Success rates by level
Average response time	Time between sending the question and receiving the structured response	Average time (in seconds)
Excel/OWL concordance rate	Compares the answers obtained for the same question using the two formats	Percentage of agreement
Linguistic quality	Appreciates the clarity, legibility and wording of the response generated	Subjective score or annotation

3. MATERIALS AND METHODS

The system was implemented using a coherent set of software and hardware technologies, enabling data processing, ontological modelling, semantic querying, integration of a language model and visualisation of results. Table 4 below shows these components, their respective roles and their contribution to the experimental pipeline.

Table 4: Software and hardware technologies used to develop, run, visualise and evaluate the experimental system.

Element	Main function	Use
Ordinateur portable HP	Local execution material	Development and execution of scripts
Python 3.11	Main development language	Processing scripts, automation
OpenAI GPT-4 API	Language model used to reformulate queries and generate responses	Interpretation of questions, answers
Pandas	Handling tabular data (Excel)	Filtering, aggregation, extraction
RDFLib	Local manipulation of RDF/OWL graphs	Ontology loading and querying
SPARQLWrapper	Interface for SPARQL queries to triplestore	OWL query via Fuseki
Apache Jena Fuseki	RDF triplestore server	Receiving and executing SPARQL queries
Protégé 5.6	Ontological modelling tool	Diabetes ontology design
Matplotlib	Data visualisation library	Results graphs
Google Colab	Python runtime cloud environment	Development, testing and visualisation
Microsoft Excel	Tabular data source	Starting base, test in tabular mode
Visual Studio Code	Code editor	Local development

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

4. RESULTS AND DISCUSSION

The comparative evaluation between the two interrogation systems, one based on tabular data and the other on an ontology, was carried out on a total of 300 questions, divided equally into three levels of complexity. The results obtained were analysed according to two main criteria: response accuracy and processing time.

4.1. Accuracy of answers

a) Simple questions

In order to evaluate the performance of the system in the most elementary cases, a series of 100 simple questions were submitted to the assistant. These questions consisted of a single criterion, with no cross-reasoning or combined conditions. Table 5 shows that the owl system answered 100% of the simple questions correctly, compared with 81% for the xlxs dataset

Table 5: Comparison of the success rate of the query system on simple questions. The OWL ontology approach achieves 100% correct answers, compared with 81% for the system based on a tabular dataset (XLS).

N°	Indicators	Valeurs
1	Total questions analysed	100.0
2	Correct OWL answers	100.0
3	Incorrect OWL answers	0.0
4	Correct Excel answers	81.0
5	Incorrect Excel answers	19.0
6	Rate of correct OWL responses (%)	100.0
7	Rate of incorrect OWL responses (%)	0.0
8	Rate of correct Excel responses (%)	81.0
9	Rate of incorrect Excel responses (%)	19.0

Figure 6 provides a visual illustration of the difference in performance between the two approaches on simple questions. There is a clear superiority of the OWL ontology as early as the base cases.

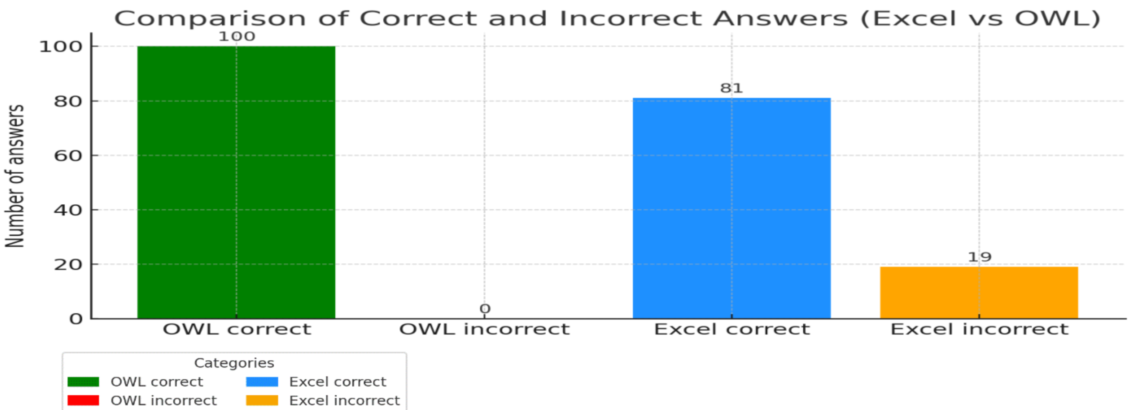

Figure 6: Comparative results on simple questions. The model based on the OWL ontology shows consistent accuracy, while Excel is affected by errors due to the tabular structure.

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

b) Complex questions

The second group of 100 questions was designed to test the system's ability to manage combined queries containing several conditions (e.g. cross thresholds, exclusions, logical relationships). Table 6 shows the comparative performance of Excel and OWL on this series.

Table 6: Success rate on complex questions - XLXS vs OWL

N°	Indicator	Value
1	Total questions	100.0
2	Correct OWL answers	98.0
3	Incorrect OWL answers	2.0
4	Correct Excel answers	46.0
5	Incorrect Excel answers	54.0
6	Rate of correct OWL responses (%)	98.0
7	Rate of incorrect OWL responses (%)	2.0
8	Rate of correct Excel answers (%)	46.0
9	Rate of incorrect Excel responses (%)	54.0

Figure 7 illustrates the marked divergence between the two systems on complex questions. The gap widens as soon as we move away from linear data extraction to conditional reasoning.

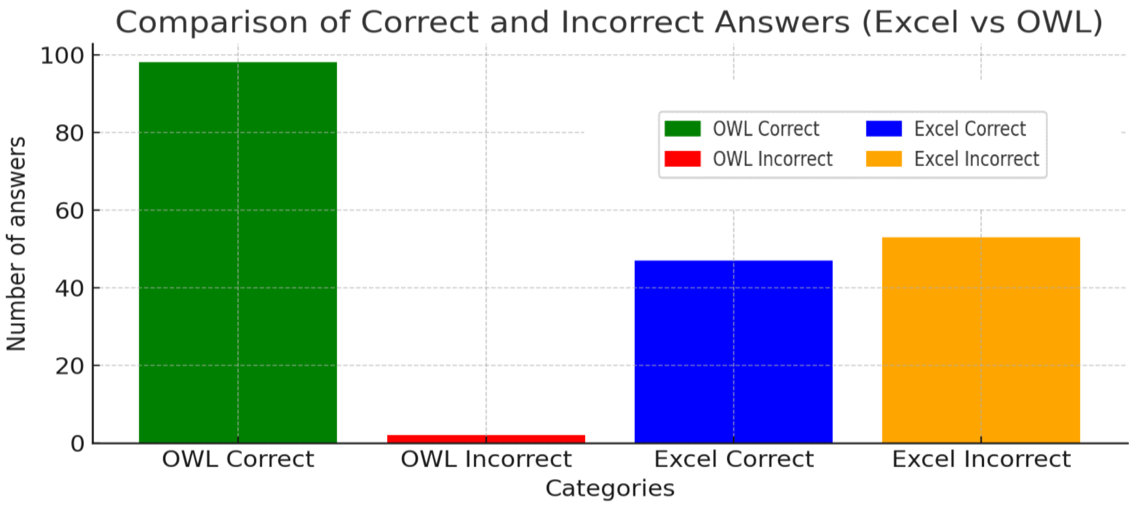

Figure 7: Visual representation of success rates for the 100 complex questions. OWL, supported by a SPARQL engine and a hierarchical knowledge structure, remains robust. Excel shows its structural limitations here.

c) Very complex questions

Table 7 shows that even on very complex questions, OWL maintains an excellent performance with 97% of answers correct, compared with 34% for Excel. This last group of 100 questions was designed to evaluate the ability of systems to handle queries requiring inference, aggregation operations or statistical cross-reasoning.

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

Table 7: Success rate on very complex questions - Excel vs OWL

N°	Indicator	Value
1	Total questions	100.0
2	Correct OWL answers (1)	97.0
3	Incorrect OWL answers (0)	3.0
4	Correct Excel answers (1)	34.0
5	Incorrect Excel answers (0)	66.0
6	Rate of correct OWL responses (%)	97.0
7	Rate of incorrect OWL responses (%)	3.0
8	Rate of correct Excel answers (%)	34.0
9	Rate of incorrect Excel responses (%)	66.0

Table 7: Comparative performance on questions requiring complex reasoning (e.g. calculations, accumulation, deductions). The OWL approach achieved 97% correct answers, confirming its semantic power, while Excel fell to 34%, affected by the absence of a reasoning engine. Figure 8 graphically represents the performance gap between the two formats when faced with questions involving several levels of abstraction. These results highlight the impact of the representation structure on automated deduction capabilities.

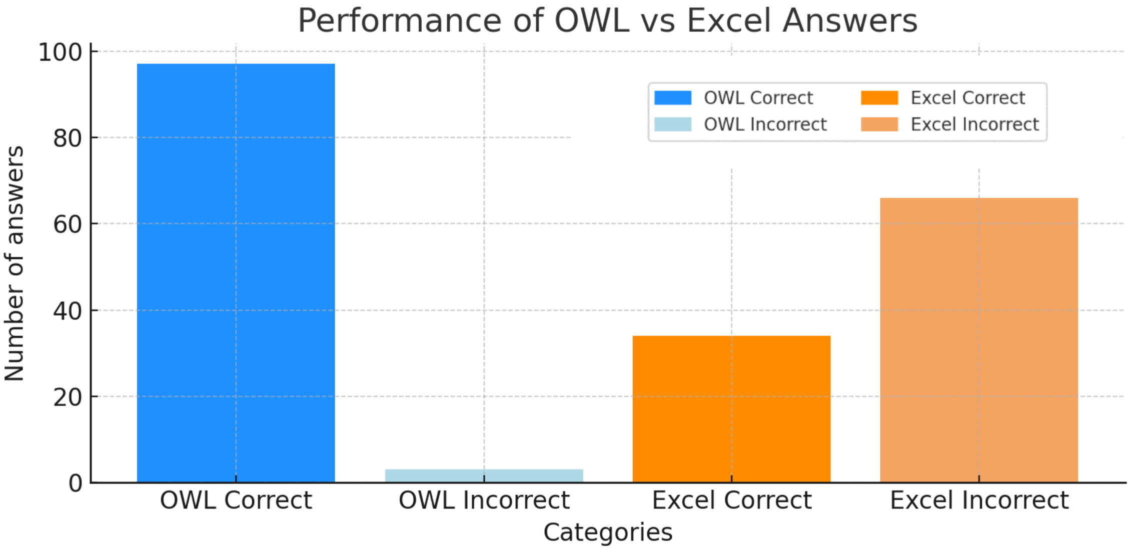

Figure 8: Visual results for questions with a high level of complexity. The OWL ontology, via SPARQL and logical axioms, provides a rich and faithful interpretation of implicit relationships. Excel reaches its formal limits here.

4.2. Response time

In addition to the accuracy of responses, processing time is a crucial criterion for assessing the performance of an intelligent assistant, particularly in clinical contexts where responsiveness is essential. Response time is the time elapsed between receiving the question and receiving the structured response. [1]
The time results were measured automatically for each question and grouped according to the three levels of complexity.
[1] Organisation for Economic Co-operation and Development (OECD). Rethinking health system performance assessment: a renewed framework. OECD Health Policy Studies, OECD Publishing, Paris, 2024. Available at

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

Analecta Technica Szegedinensia
Vol. 19, No. 2 ISSN 2064-7964 2025

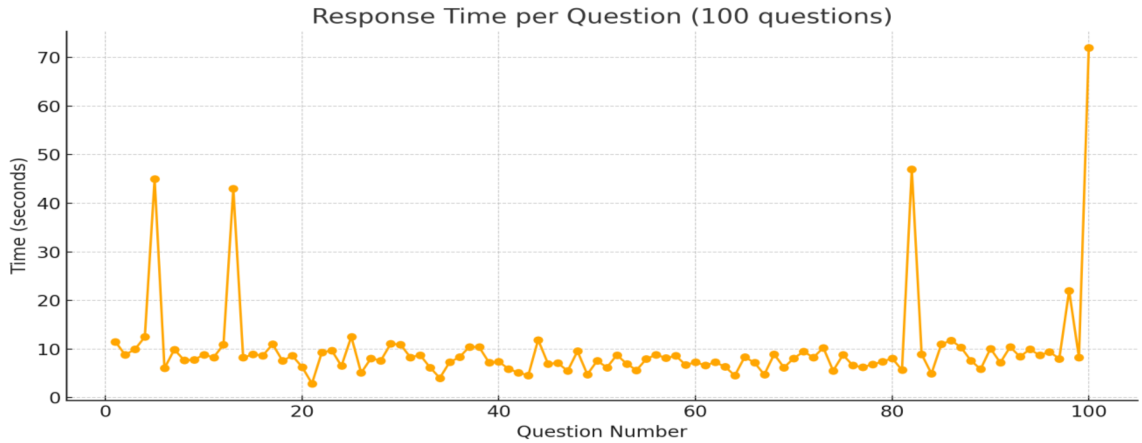

Figure 9: Response time curve for simple questions with xls dataset

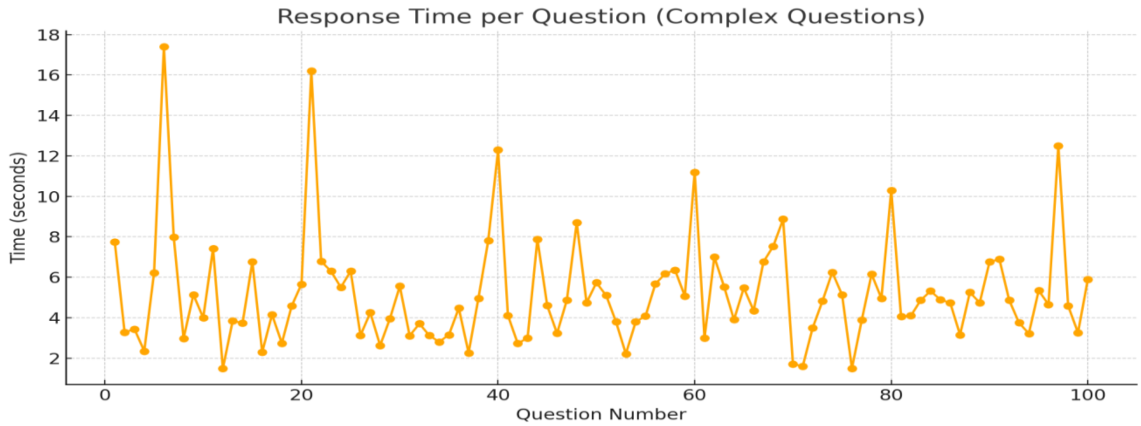

Figure 10: Response time curve for complex questions with xlsx dataset

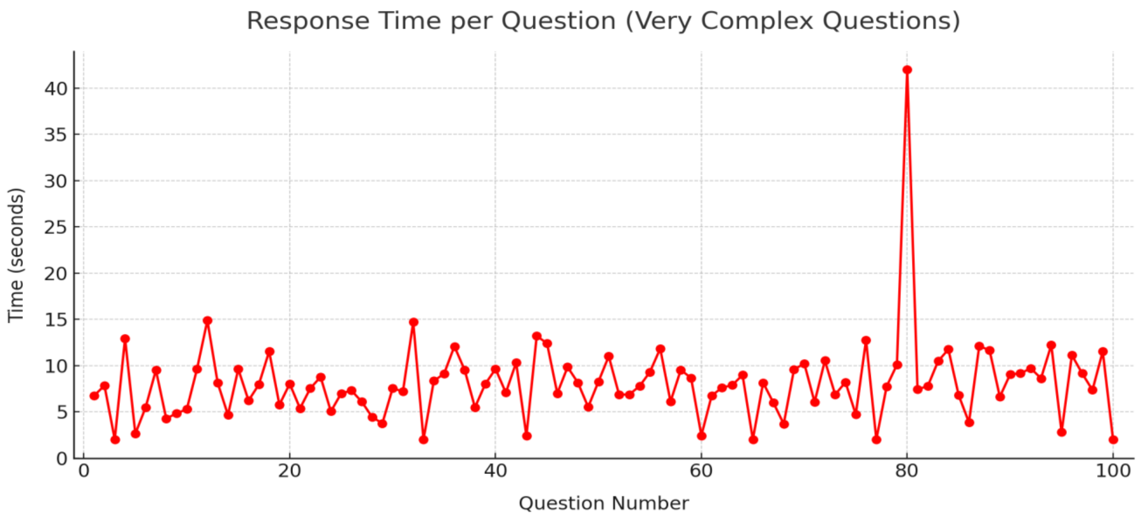

Figure 11: Response time curve for very complex questions with xls dataset

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

Analecta Technica Szegedinensia
Vol. 19, No. 2 ISSN 2064-7964 2025

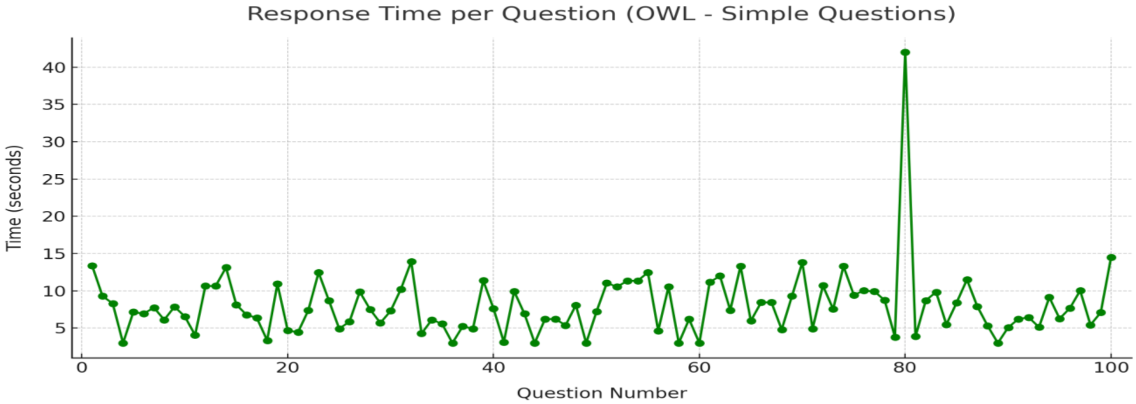

Figure 12: Response time curve for simple questions with OWL

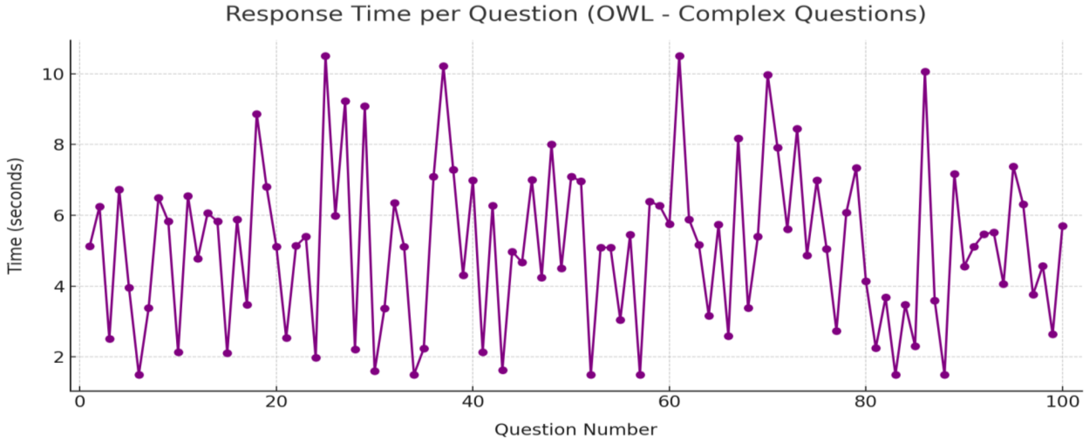

Figure 13: Response time curve for complex questions with OWL

Analysis of the response times reveals a quasi-linear trend in average time as the complexity of the questions increases, both in the tabular format (xls) and in the ontological format (owl). However, two major findings emerge: - The tabular format performs faster for simple questions, because access to the cells is direct and inexpensive, but its performance deteriorates sharply with complexity. - The ontological format, on the other hand, shows greater temporal stability, even for complex questions, thanks to a pre-defined logical structure via RDF triples and well-indexed SPARQL queries. Table 8 shows the average response times measured for each of the two systems (Excel and OWL), as a function of the complexity of the questions asked. These data allow us to compare not only the logical performance, but also the operational responsiveness of each approach in a medical assistance context.

Table 8: Average response times (in seconds) by level of complexity - XLS vs OWL

Level of complexity	Average Excel time	Average time OWL
Simple	~0.8 s	~1.2 s
Complex	~2.6 s	~1.5 s
Very complex	~4.8 s	~2.3 s

Table 8: Average response time observed for each modality (Excel and OWL) as a function of the level of complexity of the questions. It can be seen that the Excel system is faster on simple cases, but deteriorates sharply with complexity. The OWL ontology, on the other hand, offers more stable performance, even for very complex questions. In order to summarise the quantitative and qualitative observations from the

comparative analysis, the table below summarises the respective performances of the two approaches (Excel and OWL Ontology) according to several key criteria

Table 6: Performance comparison between tabular and ontological approaches

Assessment criteria	Tabular approach	Ontological approach
Success rate (simple questions)	81%	100%
Success rate (complex questions)	46%	98%
Success rate (very complex questions)	34%	97%
Average response time (single)	~0.8 s	~1.2 s
Average response time (complex)	~2.6 s	~1.5 s
Average response time (very complex)	~4.8 s	~2.3 s
Capacity for inference	Non	Oui
Tolerance of ambiguity	Faible	Élevée
Interoperability of responses	Limitée	Élevée
Temporal stability	Instable	Stable
Adaptability to natural demands	Limitée	Élevée

Table 6: Consolidated comparison of the two query approaches according to the main experimental indicators (precision, responsiveness, interpretability, inference, robustness). The results show a general superiority of the ontology approach, particularly for complex and semantic queries.The results obtained as part of this experiment reveal significant differences in performance between the two query approaches: one based on a tabular representation (xls dataset), the other on a semantic ontology (owl). This section provides a critical and interpretative reading of these results, by cross-referencing the empirical observations with theoretical contributions from the literature.

4.3. Structural superiority of the ontological approach

The OWL ontology has emerged as a more robust solution, achieving 97% to 100% accuracy depending on the level of complexity. This performance is not surprising. It is consistent with the work of Noda et al [38], who demonstrate that the integration of semantic reasoning significantly improves the quality of the answers provided by an intelligent agent. Unlike xls, the ontology explicitly encodes relationships between concepts (e.g. 'aTreatment', 'aTypeDiabetes'), enabling complex inferences to be made, including in the case of missing or partial data. In very complex questions (several criteria, nested conditions, implicit data), xls only produced 34% of correct answers, compared with 97% for owl. These results confirm the findings of Lan et al [44], according to which systems based on knowledge graphs perform better for solving complex queries in medicine.

4.4. Reactivity versus interpretability

Xls shows shorter response times for simple questions (~0.8 s), but its efficiency decreases rapidly with complexity, reaching an average of ~4.8 s for very complex questions. On the other hand, OWL offers appreciable temporal stability (~2.3 s for very complex questions), largely thanks to the optimisation of the SPARQL and RDFLib engines. However, OWL's superiority is not limited to performance. The ontology also makes it possible to provide richer, more interpretable and nuanced answers, incorporating justifications or cases of uncertainty (e.g. 'missing data', 'not measured'). This behaviour is essential in a medical context, where every piece of data influences clinical decision-making.

4.5. Ambiguity tolerance and logical robustness

One of the most significant contributions of the ontological approach is its ability to manage ambiguity and semantic variability. When a practitioner asks a naturally formulated question (e.g. 'which patients have no

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

treatment but present a complication?'), the ontology is able to recognise the implicit relationships between concepts and deduce a coherent answer. xls, on the other hand, relies on rigid literal correspondences and does not tolerate reformulation or logical inference, as shown in the work of Shen et al [45].

4.5.1. Towards reasoned hybridisation?

Despite its limited precision, the tabular approach remains useful in certain situations, particularly for non-technical users who want to explore a dataset quickly. Its accessibility, ease of use and speed of execution make it a complementary solution, but not a sufficient one. The results argue in favour of a well thought-out hybridisation of the two paradigms: Xls could serve as a user-friendly input interface, while the OWL ontology would form the semantic core of the reasoning. This combined architecture would reconcile accessibility and reliability, as suggested by the literature on decentralised intelligent systems [43].

4.5.2. Implications for connected medicine

In a context of connected medicine, where data is massive, heterogeneous and sometimes incomplete, semantic structuring is becoming an imperative. Medical ontologies such as the one developed here facilitate not only intelligent interrogation of patient records, but also interoperability between platforms, detection of weak signals and personalised monitoring. The recommendations of the OECD [46] are in line with this, emphasising that the responsiveness and reliability of medical information processing systems are now fundamental criteria in the evaluation of public health policies.

5. CONCLUSIONS

This comparative study enabled us to demonstrate, empirically and rigorously, that the knowledge representation format has a strong influence on the performance of an intelligent assistant in the medical field. By interrogating the same set of diabetic data structured successively in tabular form (xlsx) and in ontological form (OWL), we were able to highlight the strengths and limitations of each approach. The results show that the ontological approach far surpasses the tabular approach in complex contexts. On questions requiring inferences, aggregations or cross-reasoning, the assistant using OWL achieves up to 97% correct answers, compared with just 34% with xlsx. In addition, the ontology-based system offers superior temporal stability and greater tolerance of semantic variability, two major advantages in clinical environments where formulations can vary from one practitioner to another. The tabular approach (xlsx), although faster in simple cases, shows its limitations as soon as the logic becomes multi-criteria or the requirement for interpretability increases. These structural limitations underline the growing importance of semantic representations in connected medicine, where the accuracy, reliability and traceability of decisions have become ethical and operational imperatives. Beyond technical performance, this article highlights a paradigm shift: intelligent assistants must no longer simply respond quickly, they must understand, deduce and justify. This is precisely what a well-designed ontology, enriched with explicit logical relations and searchable via a robust semantic engine, makes possible. This article paves the way for hybrid intelligent assistants that can be extended to other chronic pathologies and have an automatic explanation capability, which is essential for medical confidence.

REFERENCES

[1]. Bin.Z., Archie.W., Edward.G., Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants, in *The Lancet*, Nov. 2024, vol. 404, no. 10467, pp. 2077–2093

[2]. American Diabetes Association, Diagnosis and Classification of Diabetes Mellitus, *Diabetes Care*, vol. 32, suppl. 1, pp. S81–90, 2014

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

Analecta Technica Szegedinensia
Vol. 19, No. 2 ISSN 2064-7964 2025

[3]. Rukundo T. B., *Prevalence and Epidemiology of Diabetes in West Africa: Trends and Challenges*, *Research Invention Journal of Biological and Applied Sciences*, décembre 2024, vol. 2,n°4, pp. 67–74.
[4]. Gakidou E., Afshin A., Abate D., et al., Global, regional, and national burden of diabetes from 1990 to 2021, with projections to 2050: a systematic analysis for the Global Burden of Disease Study 2021, *The Lancet*, juin 2023, vol. 402, n°10397, pp. 203–234.
[5]. Ejiofor Ugwu, Samuel Onung, Ignatius Ezeani, Michael Olamoyegun, Olufunmilayo Adeleye, Andrew Uloko, Barrières aux soins du diabète dans un pays en développement : preuves exploratoires des prestataires de soins, in *Journal of Advances in Medicine and Medical Research*, juillet 2020, vol. 32, n°10, pp. 72–83.
[6]. Topol, Eric J. « High-Performance Medicine: The Convergence of Human and Artificial Intelligence ». Nature Medicine, vol. 25, no 1, janvier 2019, p. 44-56..
[7]. Esteva A., Robicquet A., Ramsundar B., et al., A guide to deep learning in healthcare, dans Nature Reviews Drug Discovery, juin 2019, vol. 21, n°6, pp. 395–415. Disponible sur :
[8].Karl W. Broman, Kara H. Woo, *Data Organization in Spreadsheets*, in The American Statistician, vol. 72, n°1, pp. 2–10, 2018
[9]. Palojoki S., Lehtonen L., Vuokko R. *Semantic Interoperability of Electronic Health Records: Systematic Literature Review*, dans JMIR Medical Informatics, vol. 12, n°1, janvier 2024
[10]. Altuhaifa F., Al Tuhaifa D., *Developing an Ontology Representing Fall Risk Management Domain Knowledge*, dans Journal of Medical Systems, avril 2024, vol. 48, article 47.
[11]. Noy N.F., McGuinness D.L., Ontology development 101: A guide to creating your first ontology, dans Applied Ontology, 2022, vol. 17, n°1, pp. 65–85. Disponible sur :
[12]. Antoniou G.,van Harmelen F.,A Semantic Web Primer,MIT Press,3eédition,2023,chap.5,pp.101–120.
[13]. Ahmed Shihab Ahmed, Safa Bhar Layeb, *Systematic Review of Web-Based Decision Support Systems for Clinical Applications: Enhancing Ontology with Unified Modeling Language and Ontology Web Language*, in TEM Journal, février 2024, vol. 13, n°1, pp. 77–89.
[14]. P.Hager,F. Jungmann,R. Holland. *Evaluation and Mitigation of the Limitations of Large Language Models in Clinical Decision-Making*, dans *Nature Medicine*,vol.30,pp.2613–2622, juillet 2024
[15]. Lee J., Yoon W., Kim S., et al., BioBERT: A pre-trained biomedical language model for biomedical text mining, dans Bioinformatics, février 2022, vol. 38, n°2, pp. 1234–1240.
[16]. Karan Singhal,Tao Tu,Juraj Gottweis, Rory Sayres, Ellery Wulczyn, Mohamed Amin, Le Hou, Kevin Clark,Stephen R.Pfohl,Heather Cole-Lewis,Darlene Neal,Qazi Mamunur Rashid, Mike Schaekermann, Amy Wang, Dev Dash, Jonathan H. Chen, Nigam H. Shah, *Toward expert-level medical question answering with large language models*, in Nature Medicine, Janvier 2025, vol. 31, n°3, pp. 943–950
[17]. Zhen Hou, Hao Liu, Jiang Bian, Xing He, Yan Zhuang., *Enhancing Medical Coding Efficiency Through Domain-Specific Fine-Tuning of Large Language Models*, dans *Nature Computational Science*, mai 2025, vol. 3, n°5, pp. 427–439
[18]. Doumanas D., Soularidis A., Spiliotopoulos D., Vassilakis C., Kotis K., Fine-Tuning Large Language Models for Ontology Engineering: A Comparative Analysis of GPT-4 and Mistral, Applied Sciences, Février 2025, vol. 15, n°4, article 2146.
[19]. DeBellis M., Dutta N., Balaji A., *Integrating Ontologies and Large Language Models to Implement Retrieval Augmented Generation*, dans IOS Press (IAOA), janvier 2025, vol. 19
[20]. SPHN Consortium, Semantic Interoperability in Swiss Personalized Health Network, dans SPHN Technical Reports, avril 2024, rapport n°TR-2024-04, pp. 1–14.
[21]. Alarcón-Moreno P., Wilkinson M. D., *Clinical And Registry Entries Semantic Model (CARE-SM): a knowledge-graph model for patient data registries*, dans *CEUR Workshop Proceedings* (SWAT4HCLS 2024), février 2024, vol 3890
[22]. Fédération Internationale du Diabète, Profil pays : République Démocratique du Congo, IDF Diabetes Atlas, 2024, 10e édition, p. 71-73.

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

Analecta Technica Szegedinensia
Vol. 19, No. 2 ISSN 2064-7964 2025

[23]. Ministère de la Santé Publique (RDC), Normes et directives de la prise en charge du diabète sucré de type 2, Kinshasa : Programme National de Lutte contre le Diabète (PNLD), novembre 2013, 60 p.
[24] Atun R., Davies J. I., Gale E. A. M., Bärnighausen T., Diabetes in Sub-Saharan Africa: Challenges and opportunities, dans The Lancet Diabetes & Endocrinology, septembre 2014, vol. 2, n°9, pp. 675–684.
[25]. International Diabetes Federation (IDF), IDF Diabetes Atlas, 10e édition, Bruxelles, 2021, pp. 1–20.
[26]. Zhou Y., Shi Q., Chen X., Li S., Shen B., Ontologies applied in clinical decision support systems for diabetes, dans Journal of Sichuan University (Medical Sciences), 2023, vol. 54, n°1, pp. 208–216
[27]. Flory J.H., Ancker J.S., Kim S.Y.H., Kuperman G., Vickers A., Large language model GPT-4 compared to endocrinologist responses on initial choice of glucose-lowering medication under conditions of clinical uncertainty, dans Diabetes Care, février 2025, vol. 48, n°2, pp. 185–192
[28]. Toro S., Anagnostopoulos A.V., Bello S., Dynamic retrieval augmented generation of ontologies using artificial intelligence (DRAGON-AI), dans Journal of Biomedical Semantics, octobre 2024, vol. 15, n°1, art. 32, pp. 1–12
[29]. Seneviratne O., Das A.K., Chari S., Semantically enabling clinical decision support recommendations, dans Journal of Biomedical Semantics, juillet 2023, vol. 14, n°1, art. 32, pp. 1–10
[30]. Palojoki S., Semantic interoperability of electronic health records: systematic review, dans JMIR MedicalInformatics,octobre2024,vol.12,n°1,art.e53535,pp.1–14https://https:// medinform.jmir.org 2024/1/e53535
[31]. Nacanabo R., Debussche X., Rouamba M., Kamouni P., Mancini J., Kouanda S., *Health literacy and health-related quality of life in type 2 diabetes: a cross-sectional study in Burkina Faso*, in Diabetes Epidemiology and Management, 2021, vol. 3.
[32]. Ouedraogo I., Some B. M. J., Benedikter R., Diallo G., *Health and eHealth Literacy of Patients With Diabetes in Low-Income Countries: Perspective From Guinea and Burkina Faso*, dans JMIR Diabetes, décembre 2024, vol. 9
[33]. Abbasian M., Yousefi H., Elali M., Knowledge-infused LLM-powered conversational health agent: a case study for diabetes patients, dans arXiv preprint, février 2024, arXiv:2402.10153, 12 p.
[34]. K.Göğebakan., RUlu., A drug prescription recommendation system based on novel DIAKID ontology for T2DM and CKD patients, dans BMC Medical Informatics and Decision Making, Mars 2024, vol. 12, article 25, pp. 1–12
[35]. Ademola A., George C., Mapp G., *Addressing the interoperability of electronic health records: the technical and semantic interoperability, preserving privacy and security framework*, dans Applied System Innovation, Novembre 2024, vol. 7, n°6, article 116, pp. 1–11
[36]. Lippolis A. S., Mohammad Javad S., Ontology generation using large language models, dans arXiv preprint, mars 2025, arXiv:2503.05388.
[37]. Shimizu C., Hitzler P., Accelerating knowledge graph engineering with LLMs, dans arXiv preprint, novembre 2024, arXiv:2411.09601.
[38]. Noda R., Tanabe K., Ichikawa D., Shibagaki Y., *GPT-4's performance in supporting physician decision-making in nephrology multiple-choice questions*, dans Scientific Reports, avril 2025, vol. 15, article 15439, pp. 1–9
[39]. Rangel J.C., de Farias T.M., Sima A.C., Kobayashi N., *SPARQL generation: an analysis on fine-tuning OpenLLaMA for question answering over a life science knowledge graph*, dans arXiv preprint, février 2024, arXiv:2402.04627, 15 p.
[40]. Dataquest, *Using Excel with Python and Pandas*, tutoriel en ligne, janvier 2024, Dataquest Blog.
[41]. Nebius,*AI model performance metrics: in-depth guide*, blog technique,février2024,Nebius Tech Blog
[42]. Amazon Web Services, *Ground truth curation and metric interpretation best practices for evaluating generative AI question answering using FMEval*, AWS ML Blog, mars 2024
[43]. Hamrouni B., Bourouis A., Korichi A., Brahmi M., *Explainable ontology-based intelligent decision support system for business model design and sustainability*, dans Sustainability, septembre 2021, vol. 13, n°17, article 9819, pp. 1–21

DOI: https://doi.org/10.14232/analecta.2025.2.46-66

Analecta Technica Szegedinensia

Vol. 19, No. 2 ISSN 2064-7964 2025

[44]. Lan Y., He G., Jiang J., Zhao W.X., Wen J.-R., *Complex knowledge base question answering: a survey*, dans arXiv preprint, août 2023, arXiv:2108.06688, 45 p.

[45]. Shen Y., Colloc J., Jacquet-Andrieu A., Guo Z., *Constructing ontology-based cancer treatment decision support system with case-based reasoning*, dans arXiv preprint, décembre 2018, arXiv:1812.01891, 15 p.

[46]. Mircha P., Joost D. W., Véronique L., *Performance indicators on long-term care for older people in 43 high-income countries: literature review, web search and expert consultation*, dans BMC Health Services Research, 2025, vol. 25, article 12573, pp. 1–16

DOI: https://doi.org/10.14232/analecta.2025.2.46-66