# Semantic Web Technologies in Healthcare

A Scoping Review

Nelson Miranda
nelson.miranda@usp.br
Institute of Mathematical and Computer Sciences, USP
São Carlos, São Paulo, Brazil

Matheus Matos Machado
matheusmatos@usp.br
Institute Of Mathematical And Computer Sciences, USP
São Carlos, São Paulo, Brazil

Dilvan A. Moreira
dilvan@usp.br
Institute Of Mathematical And Computer Sciences, USP
São Carlos, São Paulo, Brazil

## ABSTRACT

This scoping review explores the application of Semantic Web technologies in healthcare, focusing on enhancing Electronic Health Records (EHRs). The review synthesizes research from various databases, identifying the role and impact of technologies such as RDF, OWL, and SPARQL in improving data interoperability and management within healthcare systems. Through a systematic classification and analysis of the literature, significant advancements and existing gaps in current research are highlighted. The findings suggest that while Semantic Web technologies have facilitated substantial improvements in data handling and system interoperability, challenges remain in full integration across diverse health information systems. This review underscores the potential of these technologies to transform healthcare practices by enabling more effective data integration, discovery, and management.

## KEYWORDS

semantic web technologies, Electronic Health Records, healthcare data

## 1 INTRODUCTION

The vision of the Semantic Web, as initially articulated by Berners-Lee et al. [17], promised a revolutionary transformation of the Internet into a space where data would not only connect through hyperlinks but would also describe objects, their properties, and relationships in a manner that machines could effectively understand and process. This paradigm envisioned agents that could autonomously read, interpret, and act upon web data to perform tasks on behalf of humans. Even though the full realization of the Semantic Web as a ubiquitous artifact remains elusive, the development and definition of its underlying technology stack—often visualized as a "wedding cake" model—have led to significant advancements. These technologies extend the current web by enhancing its semantic capabilities and machine-readability, marking crucial steps towards realizing interoperable systems that harness linked data for diverse applications. This article reviews the scope and impact of these technologies, particularly in how they have reshaped data interaction in healthcare.

Following the foundational perspectives established by Berners-Lee, Hendler e Lassila[17], the field of Semantic Web has been explored through various lenses. As discussed by Hitzler [32], the Semantic Web can be understood not only as an ambitious overhaul of the current World Wide Web but as a multi-faceted domain where data becomes machine-understandable by adopting sophisticated metadata schemas. This narrative has been important in promoting the development of technologies like RDF, OWL, and SPARQL, which serve as the backbone for creating, maintaining, and applying ontologies. These technologies enable data sharing, discovery, integration, and reuse beyond the confines of the Web, proving beneficial even in non-web contexts. This broader applicability hints at the Semantic Web's potential to revolutionize data handling by fostering deeper connectivity that transcends traditional hyperlinks to create a network of semantically rich and interconnected data entities.

Since the end of the DARPA DAML program in 2006, Semantic Web technologies have increasingly been applied to healthcare, significantly enhancing Electronic Health Records (EHRs). Technologies like OWL have advanced the interoperability of EHRs through ontologies such as Gene Ontology and SNOMED CT, which predate the Semantic Web's formalization and have evolved considerably. The MediBot chatbot developed by Avila et al. [16] exemplifies these technologies' practical uses in healthcare by fetching drug information and comparing prices, demonstrating the Semantic Web's broad impact on medical data management and usability across various healthcare systems.

A scoping review is defined as a type of research synthesis that provides an initial assessment of the size and scope of available research literature, identifies knowledge gaps, and clarifies concepts. According to Munn et al. [42], this approach is particularly useful when systematic reviews cannot meet the specific objectives or requirements of the knowledge users. Specifically, the primary objective of our scoping review in this context is to offer a comprehensive overview of how the technologies of the Semantic Web are applied in healthcare, particularly to improve Electronic Health Records (EHRs). It involves exploring the use of these technologies in health data management and identifying potential research gaps in existing literature. The scoping review aims to lay the groundwork for subsequent systematic investigations by defining key concepts and mapping out the extent and focus of the current research landscape.

This article is structured in a manner that gives a general overview and analysis of the application of Semantic Web technologies in healthcare. Section 2 provides the background knowledge required for the review, which is the basis for comprehending Semantic Web technologies' technical and theoretical aspects. The third section discusses related works, highlighting previous studies and existing

In: I Workshop de Revisões Sistemáticas de Literatura em Sistemas Multimídia e Web (WRSL+ 2024). Anais Estendidos do XXX Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia'2024). Juiz de Fora/MG, Brazil. Porto Alegre: Brazilian Computer Society, 2024.

ISSN 2596-1683

literature that inform and contextualize this research. The fourth section details the research methodology employed to gather and analyze the data. The fifth section reports the findings from the review, presenting an empirical evaluation of how these technologies are currently utilized in healthcare settings. The sixth section discusses such findings by interpretation and brings out the implications of this for future technology development and application within the context of healthcare. The last section draws conclusions from the study, summarizes key insights, and suggests directions for further research.

## 2 BACKGROUND

This section defines key terms and assets foundational to our scoping review, particularly focusing on Semantic Web technologies in healthcare. By clarifying these concepts, we establish a unified understanding and ensure consistent communication throughout the review.

EHRs include a broad, shareable collection of a patient's records across many healthcare settings, such as medical history, medication, allergies, laboratory results, and others. Sharing important information about a patient in this manner becomes easy with EHRs; hence, the delivery of health services becomes efficient and effective for one to view the patient's general health status. The technology supports better decision-making to increase coordination among healthcare providers, thereby improving overall outcomes. They support individual and population health management, enhancing coordinated care and patient outcomes[2, 11].

In contrast, Electronic Medical Records (EMRs) are confined to specific healthcare provider encounters and do not typically transfer between systems. They focus on documenting clinical and administrative data for use within individual practices[11].

Health Information Systems include EHRs, EMRs, and those systems that manage the billing, scheduling, and patient management of information. HIS works to improve healthcare efficiency and quality by providing various data sources so that decisions can be made through integration[51].

RDF (Resource Description Framework), RDF Schema (RDF-S), and OWL (Web Ontology Language) are critical components of the Semantic Web stack, which facilitate advanced data representation and interoperability. The primary objective of RDF is to deliver a generic model for describing the meaning or semantics of information resources on the web. This structure is key in integrating diverse data sources, enhancing data merging capabilities regardless of differing underlying schemas[4, 6].

Expanding on RDF, RDF-S offers a semantic vocabulary that helps define and categorize properties and classes of RDF resources, such as establishing class hierarchies that increase the complexity and usability of the data models. For example, RDF-S allows for the specification of class relationships, making it possible to express that a Dog is a type of Animal, thus linking the classes Dog and Animal in a meaningful way[3].

Building upon RDF and RDF-S, OWL provides a more expressive framework for detailed ontology development. It supports richer descriptions and reasoning about the data, incorporating advanced features like class equivalence, property characteristics, and complex class hierarchies. OWL's capabilities are crucial for applications that require deep knowledge representation and reasoning, such as those in artificial intelligence and semantic web services[3, 6].

SWRL (Semantic Web Rule Language) enhances OWL (Web Ontology Language) by integrating rule-based reasoning capabilities through its combination with RuleML. Introduced by the W3C in 2004, SWRL extends OWL's ontology frameworks with Datalog-style rules, enabling complex reasoning and knowledge inference[33]. This powerful language allows for the expression of rules that facilitate dynamic inference, making it particularly suitable for applications requiring sophisticated data analysis and manipulation.

Rules in SWRL are expressed as implications, where specific conditions lead to certain conclusions, thereby supporting applications that need to derive new knowledge from existing data under predefined conditions[33, 47]. Its integration with OWL ontologies allows for leveraging structured knowledge representation alongside dynamic rule-based processes, making SWRL an important tool for advancing capabilities in the Semantic Web.

SWRL is particularly valuable in scenarios requiring complex inference and dynamic knowledge representation, where the knowledge base needs to evolve based on new rules or information. This makes it a key tool for developers and researchers working within the Semantic Web framework, providing a mechanism for detailed and extensible knowledge management [47].

Therefore, RDF, RDF-S, OWL, and SWRL together make a strong infrastructure for the Semantic Web. This infrastructure allows one to create interlinked data structures, ensuring semantic interoperability across applications. It also allows more effective data sharing and use between systems in areas such as health, finance, and education. These technologies have advanced capabilities to support complex data relationships and optimize the overall efficiency and effectiveness of information management.

SPARQL is the query language designed for RDF data, so it represents one of the most important tools with respect to dealing with and querying data in the context of the Semantic Web framework. SPARQL allows one to easily query diverse, interlinked datasets, find various information, and derive new knowledge from these data. Its capability to handle difficult queries and output accurate results is beneficial for applications that manipulate semantically integrated and analyzed data. Developed by the W3C and officially recommended in 2008, SPARQL enables the manipulation of data stored in RDF format and supports a variety of operations, from basic data retrieval to complex queries across multiple datasets[57]. Its ability to integrate and query data from diverse sources makes SPARQL indispensable for applications requiring sophisticated data management, such as knowledge graphs and semantic searches[1].

Beyond mere data retrieval, SPARQL's role extends to supporting advanced search capabilities that consider the semantic relationships inherent in RDF data. This functionality is essential in fields like artificial intelligence, where understanding data relationships is key, and business analytics, where insights are drawn from complex interconnected data sets[27].

In summary, SPARQL facilitates robust data integration and quality management across heterogeneous sources and leverages Semantic Web technologies for more dynamic and context-aware data handling.

# 3 RELATED WORKS

The Semantic Web provides promising solutions for integrating and managing healthcare data, addressing challenges like interoperability, knowledge sharing, and decision support across diverse systems[31, 43]. By enhancing context-based information searching and integration, Semantic technologies significantly contribute to medical research[34]. These technologies also bolster Electronic Health Record (EHR) systems, improving healthcare quality through context-aware searches and rule creation[34]. However, strong security mechanisms are essential when handling sensitive health data, with various strategies developed for authentication, authorization, privacy, and other security aspects in Semantic Web applications[21]. The use of Semantic Web technologies is on the rise in healthcare and life sciences, particularly in disease-causal gene analysis, drug efficacy assessment, and building knowledge bases for biomedical research[19]. Despite some challenges, these technologies hold promise for managing healthcare big data and extracting valuable insights[30].

The systematic review by Haque et al. [31] explores the use of Semantic Web (SW) technology in healthcare, focusing on its potential to address data management challenges like information exchange, interoperability, and decision support. Analyzing 65 papers, the authors identify five key themes: e-service, disease, information management, frontier technology, and regulatory conditions. The study highlights how SW technology aids in developing e-healthcare systems that support decision-making for medical practitioners and provide patients with vital information and automated services. The review also emphasizes the importance of SW in improving knowledge exchange and data interoperability in healthcare, while noting research gaps and proposing future directions for advancing SW-based medical systems.

The semantic web is an emerging technology that improves information representation and connectivity, enabling AI-driven applications[43]. Central to this technology are ontologies, which allow for standardized sharing and reuse of concepts across different data sources. The paper highlights the growing use of semantic web technologies in healthcare, virtual communities, and information retrieval systems, emphasizing the need for appropriate ontologies to resolve ambiguities and ensure accurate interpretation of textual content. Narayanasamy et al. [43] provide a detailed review of semantic web applications in these fields, underscoring the technology's potential to transform them. As the world advances into the fourth industrial revolution, the semantic web's importance in connecting users, generating content, and enhancing computer understanding of information becomes increasingly critical[43].

This scoping review by Costa Lima et al. [21] explored security approaches for managing electronic health data using Semantic Web technologies. Analyzing 26 studies, the authors identified security mechanisms addressing key attributes like authentication, authorization, confidentiality, and privacy. These mechanisms support frameworks for access control and privacy compliance in healthcare. The review emphasizes the growing use of Semantic Web technologies in healthcare and points out the need to understand technical requirements to mitigate risks in managing personal health information, contributing to secure health information system integration.

In our scoping review, we explore how Semantic Web technologies like RDF, OWL, and SPARQL enhance data interoperability and management in healthcare, focusing specifically on Electronic Health Records (EHRs). Compared to broader studies by Haque et al. [31] and Narayanasamy et al. [43], our work is more specialized, emphasizing the technological impact on EHRs and identifying research gaps. While all reviews highlight the transformative potential of these technologies in healthcare, our analysis provides a more focused view on their practical integration into existing health systems, suggesting future directions like real-time data processing and dynamic ontology management.

# 4 MATERIALS AND METHODS

Scoping reviews are most efficient when it comes to identifying the extent of the literature that deals with a particular subject, and this helps to understand the amount of information available for research. Such reviews are applicable in summarizing information on a wide range of data and identifying research gaps. The scoping review methodology is adopted with a structured approach following the work of Arksey and O'Malley [14], further developed in guidelines provided by the Joanna Briggs Institute [39], and later by Munn et al. [42] and Peters et al. [45]. Such guidelines propose a five-stage framework for conducting the review process, as elaborated in the following subsections, to ensure the completeness of the procedure.

## 4.1 Stage 1: Identifying the Research Question

The initial stage of our scoping review aimed to systematically explore integrating Semantic Web technologies with Electronic Health Records. The research questions designed to guide this inquiry were as follows:

(1) **Which Semantic Web technologies are commonly applied in conjunction with Electronic Health Records?** The question identifies and describes the main tools and frameworks of the Semantic Web that are currently being used to upgrade the functionality and interoperability of EHRs. These tools and frameworks are described to see how they contribute to the health system's betterment of data management, sharing, and integration.

(2) **What data management perspectives are addressed by these technologies?** This involves how Semantic Web technologies participate in health data management practices related to sharing, discovery, integration, and reuse. An analysis of these aspects will determine the role of technologies in enhancing effectiveness and efficiency in healthcare data management.

(3) **What are the existing knowledge gaps regarding handling health data through Semantic Web technologies?** Identification of where research is needed, especially the limitations and challenges healthcare providers face in adopting these technologies. This includes the identification of specific barriers to implementation, both technical and organizational, and how these can affect overall uptake and impact the effectiveness of Semantic Web technologies in healthcare.

So, these questions structured the approach to gathering and analyzing literature for the review, with a view to ensuring there was a full understanding of the state of Semantic Web technologies in healthcare data management. By addressing these questions, the review systematically covered different facets of applications of the Semantic Web to point out their progress, the challenges they faced, and the opportunities for further research.

## 4.2 Stage 2: Identifying Relevant Studies

A thorough search strategy was executed to gather relevant literature from major databases, including IEEE Xplore Digital Library, Scopus, Embase, PubMed/MEDLINE, and Web of Science. The search string employed was:

```
("Semantic Web" OR "Web Ontology Language" OR "OWL" OR
"SPARQL" OR "RDF" OR "SWRL") AND ("Electronic Health
Records" OR "EHR" OR "electronic medical records" OR
"EMR" OR "health information systems")
```

This query was designed to capture all pertinent studies discussing the integration of Semantic Web technologies with various health information systems.

**Eligibility Criteria.** Inclusion criteria included reports from research studies, literature reviews, full conference papers, and non-research content, including editorials and letters if they were published in English between January 2000 and July 2024. Studies were excluded if they did not include the text words "Electronic Health Records," "semantic web," and at least one more related term in the title or abstract or only described theoretical frameworks/models without any implemented model applied to Semantic Web.

**Forward and Backward Searching.** To ensure comprehensiveness, the reference lists of each selected article were hand-searched to identify additional studies not caught in the initial database search for inclusion. We followed references of articles by using the forward and backward search methods to include appropriate studies that met these selection criteria.

## 4.3 Stage 3: Study Selection

**Article Screening and Review Process.** Three researchers and an AI-driven robot based on the GPT-4 Turbo model independently screened each retrieved article by title and abstract for eligibility. After this initial screening, the entire text of these articles was obtained for a more comprehensive review. In-depth discussions within the research team resolved any reviewer discrepancies. It is important to note that during the review process, the reviewers were aware of the journal title, authors, and their affiliated institutions.

This scoping review is not designed to strictly adhere that of a systematic-review format, however guidelines from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) have been used to ensure transparency with study selection[41].

## 4.4 Stage 4: Charting the Data

**Data Extraction Strategy.** A planned data extraction method from selected studies was utilized to obtain relevant information, ensuring that collected data adequately responded to the research questions specified in Stage 1. A breakdown of precise data pulled from each investigation is shown in the table below:

**Table 1: Data extraction strategy**

| Scope | Data to be Extracted |
|---|---|
| Summary | Title, authors, publication type, year of publication, periodic/journal, aims/objectives |
| Q1 | List of Semantic Web technologies used |
| Q2 | Data management perspectives addressed |
| Q3 | Proposed method, approach, or technology; advantages and drawbacks. |

This strategic approach ensured that all relevant facets of each article were thoroughly analyzed, from basic bibliographic information to specific details concerning the implementation and impact of Semantic Web technologies in healthcare.

## 4.5 Stage 5: Collating, Summarizing, and Reporting the Results

**Data Synthesis and Analysis.** In line with the objectives of scoping studies to provide a comprehensive overview of the reviewed literature, as described by Arksey and O'Malley [14], data from Stage 4 had been systematically organized. The classification was meticulously done and presented in tables categorized by the first author of the publication and by the semantic web technologies utilized. This organization helped visually summarize the data trends over time and across different research themes.

Other tables present the methodology, approach, or technology utilized in each study, along with their advantages and disadvantages. It includes a narrative synthesis that aims to delineate and analyze the similarities and differences observed among the studies. This synthesis explores emerging patterns, themes, and relationships, critically examining the data. The goal is to identify coherent strands of evidence and areas requiring further investigation.

# 5 RESULTS

This chapter details the process and outcomes of the scoping review conducted. Initially, a comprehensive search across selected databases yielded a total of 313 articles, with an additional 10 articles identified from other sources, summing up to 323 articles. After removing 147 duplicates, 176 articles remained for further screening. The distribution of articles by year and database can be seen in Table 2.

Upon screening titles and abstracts, the number was further reduced to 127. Subsequent full-text assessments for eligibility led to the exclusion of 32 articles due to accessibility issues (lacking free or institutional access). After applying eligibility criteria rigorously, 25 articles met all the requirements and were included in the final review. The list of accepted articles can be seen in the table 3:

The PRISMA flow diagram, presented in Figure 1, visualizes the sequential stages of the review process. This diagram provides a detailed breakdown of each step, from identifying records to the final inclusion in the review.

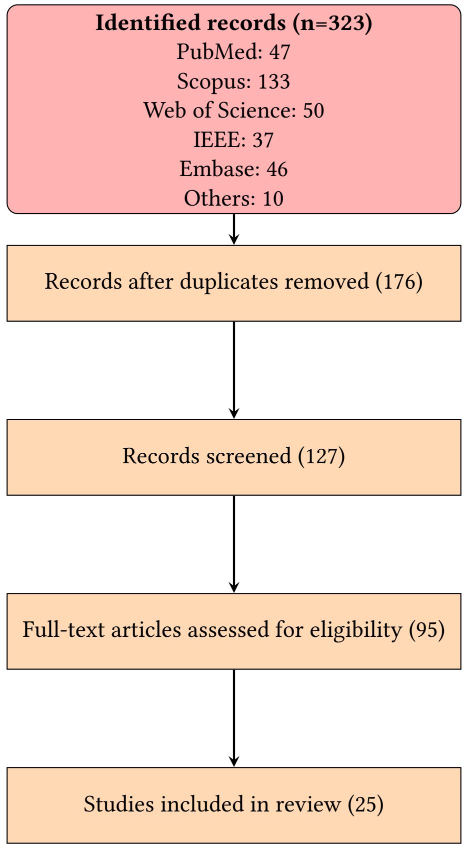


**Figure 1: PRISMA Flow Diagram of the study selection process.**

**Table 2: Publication Counts by Source and Year**

| Source | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 |
|---|---|---|---|---|---|---|
| Web of Science | 7 | 11 | 10 | 11 | 8 | 3 |
| IEEE | - | 10 | 11 | 6 | 9 | 1 |
| PubMed | - | 7 | 9 | 12 | 11 | 8 |
| Scopus | - | 26 | 38 | 31 | 26 | 12 |
| Embase | - | 6 | 6 | 18 | 10 | 6 |
| Total | 7 | 60 | 74 | 78 | 64 | 30 |

## 5.1 Which Semantic Web technologies are commonly applied in conjunction with Electronic Health Records?

In the realm of healthcare informatics, integrating Semantic Web technologies with Electronic Health Records (EHRs) has emerged as a pivotal strategy to enhance data interoperability and functionality. This subsection delves into the prevalent Semantic Web tools and frameworks that have been identified from the comprehensive review of the selected articles. These technologies are key in advancing the capabilities of EHR systems, thereby facilitating more effective data sharing and integration across diverse healthcare platforms. The findings, presented in Table 4, highlight the specific technologies employed and the corresponding studies that utilize these advancements in the context of EHRs.

**Table 4: Semantic Web Technologies (SWT) Utilized in EHRs**

| SWT | Reference |
|---|---|
| RDF/RDF-S | [9, 10, 13, 18, 20, 23, 26, 28, 37, 46, 48, 50, 52–56] |
| OWL | [7–10, 13, 20, 25, 26, 28, 35, 46, 48, 50, 52–55] |
| SPARQL | [7, 10, 13, 20, 25, 26, 28, 35, 37, 44, 46, 48, 53, 55] |
| Prova | [37] |
| SHACL | [37] |
| SWRL | [9, 28] |
| N3 | [50] |
| EYE | [50] |
| SPIN | [35] |
| Linked Data | [13, 20, 35] |
| Knowledge Graph | [53, 54] |

In the scope of healthcare, the analysis reveals a significant reliance on foundational Semantic Web technologies such as RDF, OWL, and SPARQL across various studies. These technologies, constituting the core layers of the Semantic Web stack, have been predominantly utilized to enhance the functionality and interoperability of Electronic Health Records (EHRs). The widespread adoption of RDF, OWL, and SPARQL underscores their robustness in

**Table 3: Summarized Article References**

| Ref | Title | First Author | Year | Publisher |
|---|---|---|---|---|
| [8] | Semantic architecture for interoperability... | Adel et al. | 2022 | IEEE |
| [13] | Using a personal health library... | Ammar et al. | 2021 | JMIR Publications Inc. |
| [10] | Capturing semantic relationships... | Aldughayfiq et al. | 2023 | MDPI |
| [9] | Ontological framework for standardizing... | Alahmar et al. | 2020 | Elsevier |
| [24] | Semantically rich access control... | Dixit et al. | 2022 | IEEE |
| [23] | EMR2vec: Bridging the gap... | Dhayne et al. | 2021 | Elsevier |
| [25] | Structuring, reuse and analysis... | Duncan et al. | 2020 | Springer |
| [26] | An ontology-based Approach for... | Frid et al. | 2023 | JMIR Publications Inc. |
| [28] | A drug prescription recommendation system... | Gögebakan et al. | 2024 | Springer |
| [35] | Supporting integrated care... | Kilintzis et al. | 2019 | Elsevier |
| [36] | Uncertainty-aware text-to-program... | Kim et al. | 2022 | PMLR |
| [40] | Advanced Data Processing of Pancreatic Cancer... | Manias et al. | 2024 | MDPI |
| [37] | Modeling medical guidelines... | Kober et al. | 2022 | IOS Press |
| [44] | Knowledge graph-based question answering... | Park et al. | 2021 | PMLR |
| [46] | Cohort Identification Using Semantic... | Pfaff et al. | 2021 | Cold Spring Harbor Laboratory Press |
| [52] | Secure cloud ehr with semantic... | Walid et al. | 2021 | IEEE |
| [53] | Leveraging semantic context to... | Walid et al. | 2024 | Elsevier |
| [48] | EHR-oriented knowledge graph system... | Shang et al. | 2021 | IEEE |
| [20] | Colorectal cancer health and care... | Choudhury et al. | 2024 | Springer |
| [18] | An ontology-based system for... | Casey et al. | 2022 | IEEE |
| [56] | Leveraging genetic reports... | Zong et al. | 2021 | JMIR Publications Inc. |
| [54] | FHIR-Ontop-OMOP: Building clinical... | Xiao et al. | 2022 | Elsevier |
| [7] | Developing an Exercise Games Ontology... | Abdullah et al. | 2022 | IEEE |
| [55] | PO2RDF: representation of real-world... | Zhao et al. | 2022 | Springer |
| [50] | Predicting future state for... | Sun et al. | 2021 | Elsevier |

facilitating complex data integration and semantic querying processes, which are crucial for the efficient management of medical data and knowledge discovery in healthcare environments. Such consistent use across diverse research [7–10, 13, 18, 20, 23, 25, 26, 28, 35, 37, 44, 46, 48, 50, 52–56] indicates their effectiveness and the potential for further explorations in semantic applications within the healthcare sector.

The study by Sun et al. [50] highlights the integration of N3 logic and the EYE reasoning engine in Electronic Health Records (EHRs) to enable dynamic clinical pathway management. This approach enhances the adaptability of clinical decisions in real-time and underscores the potential of Semantic Web technologies to revolutionize patient care by providing context-aware adjustments. The use of N3 logic allows for more expressive reasoning capabilities beyond traditional RDF and OWL[15], while the EYE reasoning engine supports efficient real-time processing[22]. This methodology is particularly effective in complex healthcare settings where standard protocols may fall short, offering a pathway towards more personalized and effective healthcare solutions.

The study by Kober et al. [37] demonstrates the utility of integrating Prova and SHACL with Electronic Health Records (EHRs) to enhance clinical guideline management. Prova, a rule-based language, is adept at complex data integration and facilitates dynamic medical decision-making through its compatibility with SPARQL queries[38]. SHACL, used for data validation in RDF frameworks, ensures data accuracy, which is crucial for effective clinical decisions[5]. This integration supports real-time adaptations of medical guidelines, enhancing clinical workflows with technologies like the ABCDE assessment method[49], underscoring the significance of Semantic Web technologies in healthcare.

The concepts of Linked Data and Knowledge Graphs represent subsets of Semantic Web technologies, offering sophisticated structures for data management and integration[32]. The papers by Kilintzis et al.[35], Ammar et al.[13], Choudhury et al.[20], Walid et al.[53], and Xiao et al.[54] provide compelling evidence of the applicability of these technologies in enhancing the functionality and interoperability of EHR systems. These studies underscore the transformative potential of Semantic Web technologies in healthcare, supporting more effective data sharing, discovery, integration, and reuse. Such advancements highlight a paradigm shift towards more connected and accessible health information systems.

We categorized the tools and utilities used in the selected articles to implement Semantic Web technologies, providing a structured overview and emphasizing their roles in data interoperability and management. Table 5 classifies these tools into categories such as graph databases, ontology tools, and data conversion utilities, aiding in understanding the technological landscape and serving as a reference for researchers and practitioners. Table 5 is the revised table with the mentioned tools and utilities' URLs.

Several health standards are being enforced in the fields of Semantic Web technologies used in healthcare to improve interoperability, accessibility, and usability for EHRs. As part of our scoping

**Table 5: Implement Semantic Web Technologies**

| Category | Tools & Utilities |
|---|---|
| Graph Databases & Triplestores | GraphDB, Neo4j, Blazegraph, Apache Jena TDB2, Virtuoso |
| Ontology Tools | Protégé, Cellfie, LinkEHR |
| Libraries & Frameworks | Apache Jena, RDFlib, LDflex |
| Data Management & Conversion | OpenRefine, Ontotext Refine, D2RQ |

**Links:**
GraphDB: https://www.ontotext.com/products/graphdb/
Neo4j: https://neo4j.com/
Blazegraph: https://blazegraph.com/
Apache Jena TDB2: https://jena.apache.org/documentation/tdb/
Virtuoso: https://virtuoso.openlinksw.com/
Protégé: https://protege.stanford.edu/
Cellfie: https://protegewiki.stanford.edu/wiki/Cellfie
LinkEHR: http://www.linkehr.com/
Apache Jena: https://jena.apache.org/
RDFlib: https://rdflib.dev/
LDflex: https://github.com/solid/query-ldflex
OpenRefine: https://openrefine.org/
Ontotext Refine: https://ontotext.com/products/ontorefine/
D2RQ: http://d2rq.org/

review, we have identified and categorized the key health standards commonly integrated with Semantic Web technologies to improve data exchange, representation, and analysis across different healthcare systems. These standards were organized into categories depending on the main purpose and usage (see Table 6).

This review hopes to clarify the landscape of health standards in the context of Semantic Web applications (by mapping and classifying these standards) and serve as a reference for researchers or practitioners working at the intersection between technology and healthcare. Integration of these standards with Semantic Web technologies is necessary to develop advanced health information systems that are capable, interoperable, and effective.

**Table 6: Health Standards in Semantic Web Applications**

| Category | Standards |
|---|---|
| Interoperability Standards | HL7 FHIR, DICOM, IHE |
| Coding and Classification Systems | SNOMED CT, LOINC, ICD, ATC, NDC, RxNorm |
| Data Models and Frameworks | EN/ISO 13606, OMOP CDM, openEHR |

**Links:**
HL7 FHIR: https://www.hl7.org/fhir/
DICOM: https://www.dicomstandard.org/
IHE: https://www.ihe.net/
SNOMED CT: http://www.snomed.org/snomed-ct/
LOINC: https://loinc.org/
ICD: https://www.who.int/classifications/icd/en/
ATC: https://www.whocc.no/atc_ddd_index/
NDC: https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory
RxNorm: https://www.nlm.nih.gov/research/umls/rxnorm/index.html
EN/ISO 13606: https://www.iso.org/standard/62303.html
OMOP CDM: https://www.ohdsi.org/data-standardization/the-common-data-model/
openEHR: https://www.openehr.org/

HL7 FHIR is a compulsory standard in the United States of America and has already received acceptance into Brazil's SUS, emphasizing on the use of Semantic Web technologies for enhanced data exchange and interoperability between systems. This provides vital support to the process of modernising health information systems, improving communication among health providers and promoting patient care. In today's digital healthcare context, it cannot be overemphasized that FHIR guarantees reliable and secure information exchange among healthcare systems.

## 5.2 What data management perspectives are addressed by these technologies?

Semantic Web technologies enhance healthcare data management through several key functions. We have categorized the reviewed studies into three main groups based on their focus, noting that these categories are not mutually exclusive; a single study may span multiple groups: **Interoperability and Data Sharing** they enable semantic interoperability among different healthcare systems, facilitating better data and knowledge exchange among healthcare stakeholders, which is crucial for coordinated care (8 articles); **Data Integration and Discovery** these technologies integrate and discover medical data from varied sources, allowing efficient access to diverse health datasets and improving service delivery by linking information across platforms (25 articles); and **Data Reuse and Knowledge Management** they improve information management and clinical decision-making by enabling semantic and pragmatic levels of service and knowledge sharing, incorporating AI and IoT for enhanced diagnostic and security measures (15 articles).

*5.2.1 Interoperability and Data Sharing.* In recent times, there has been an increase in healthcare informatics, which emphasizes the significance of Semantic Web technologies in improving interoperability and data sharing within EHR systems. Studies conducted by Redwan Walid et al. (2024)[53], Vassilis Kilintzis et al. (2019)[35], and Nariman Ammar et al.[13] discuss how these technologies such as ontologies and knowledge graphs can be incorporated in such a way that EHRs become more secure, efficient and user-controlled. These studies demonstrate that by leveraging attribute-based and searchable encryption, EHR systems can achieve fine-grained access control and secure data querying, ensuring privacy and robust data management.

Furthermore, initiatives like the Personal Health Libraries[12] and the adoption of standards such as SNOMED CT, HL7 FHIR, and the OMOP Common Data Model are crucial for facilitating seamless data exchange and enhancing analytical capabilities across healthcare systems. The exploration of federated approaches, such as the Personal Health Train by Ananya Choudhury et al. (2024)[20], also underscores the potential for preserving privacy while computing quality indicators. These technologies bolster the infrastructure of health data management and pave the way for advanced applications in integrated care and artificial intelligence in healthcare.

*5.2.2 Data Integration and Discovery.* This subsection of the scoping review highlights research that has utilized the approach of Data Integration and Discovery in healthcare. It examines how recent studies have focused on improving clinical decision support systems and integrating diverse healthcare data sources. For instance, Santiago Frid et al. (2022)[26] discuss how patient data can be consolidated using ontologies and common data models to streamline healthcare processes. Similarly, Hong Sun et al. (2020)[50] and George Manias et al. (2024)[40] have developed methodologies for predictive analytics and holistic health records, enhancing clinical pathways and patient care.

Efforts such as those by Gerhard Kober et al. (2022)[37] and William D Duncan et al. (2020)[25] have applied Semantic Web technologies to model medical guidelines and structure dental data, demonstrating the significant impact of data integration. Nansu Zong et al. (2020)[56] have utilized machine learning to enhance diagnostic accuracy by integrating genetic reports with Electronic Health Records. Additionally, Zul Hilmi Abdullah et al. (2022)[7] and Redwan Walid et al. (2024)[53] have contributed to security enhancements and standardized ontology development, further illustrating the diverse applications of Semantic Web technologies in healthcare data integration and discovery. These collective efforts underscore a significant shift towards a more integrated and informed healthcare system, leveraging advanced data management strategies to improve clinical outcomes and patient care.

Several studies have explored the use of Resource Description Framework (RDF) triplestores and knowledge graphs for complex computable phenotyping and cohort identification (Pfaff et al., 2021; Shang et al., 2021)[46, 48]. Personal Health Libraries (PHLs) have been proposed to give patients more control over their digital health data and enable integration with other knowledge sources (Ammar et al. [13]). Ontology-based frameworks have been developed for standardizing and digitizing clinical pathways (Alahmar et al., 2020)[9] and integrating diverse healthcare data sources (Kilintzis et al., 2019)[35]. These approaches aim to enhance healthcare systems' interoperability, privacy preservation, and data utilization. Additionally, federated approaches using FAIR data principles[29] and the Personal Health Train infrastructure have been explored for calculating quality indicators across multiple hospitals (Choudhury et al., 2024)[20]. Overall, these studies demonstrate the potential of semantic web technologies to improve healthcare data management and decision-making.

Numerous studies have devised methods for secure, fine-grained access control to cloud-based EHRs using attribute-based encryption and ontology reasoning (Redwan Walid et al., 2024; Redwan Walid et al., 2021)[52, 53]. Graph-based representations of EHR data have shown advantages over traditional relational databases for question-answering tasks (Junwoo Park et al., 2020)[44]. Researchers have also explored methods for integrating patient data with clinical trials (Houssein Dhayne et al., 2021)[23] and transforming heterogeneous medical information into interoperable formats using semantic technologies. Efforts have been made to create clinical knowledge graphs compliant with healthcare standards like FHIR and OMOP (Guohui Xiao et al., 2022)[54], represent real-world oncology data using RDF (Yiqing Zhao et al., 2021)[55], and develop ontology-based models for achieving interoperability between distributed EHR systems (Ebtsam Adel et al., 2022)[8].

*5.2.3 Data Reuse and Knowledge Management.* Data reuse and efficient knowledge management are cornerstones in enhancing healthcare data systems. A recent emphasis on refining the functionality of EHR systems has been motivated by the power of Semantic Web Technologies. Notably, research has helped develop ontologies that standardize and combine various medical data, which leads to improved interoperability across different healthcare systems as well as coordinated care. For example, various works by Duncan et al. (2020)[25], Kilintzis et al.(2019)[35], and Alahmar et al.(2020)[9] have relied on such ontologies. Furthermore, knowledge graphs have made it possible for EHR functionality to be enriched through secure access control as discussed by Redwan Walid et al.(2024)[53] or through effective use of unused clinical information (Shang et al., 2021)[48]. Furthermore, these technologies have facilitated the representation of real-world data for precise applications such as oncology (Zhao et al., 2021)[55] and have been proven advantageous in question-answering systems within EHRs, surpassing traditional methods (Park et al., 2020)[44]. Concurrent efforts mean that information repetition plays a significant part in improving medical attention quality, helping to make choices based on examination results, and allowing general patient investigations within many areas of medicine.

Expanding on the significant function of data reuse and knowledge management for health care, recent investigations have also looked into how Semantic Web technologies can upgrade EHR systems. The ongoing research focuses on using ontologies and knowledge graphs to standardize and integrate various data sources, greatly facilitating efficient data analysis and decision-making. Notable efforts include the work by Santiago Frid et al. (2022)[26] and Shinead Casey et al. (2022)[18], which underscore the practical application of these technologies in enhancing data management systems.

Additionally, the concept of Personal Health Libraries (PHLs) has been proposed, as seen in the studies by Nariman Ammar et al.[13], to empower patients with greater control over their health data, supporting the self-management of chronic conditions. This patient-centric approach aligns with the broader goals of Semantic Web technologies to foster personalized and patient-driven healthcare.

Ontology-based systems have also been tailored for specific healthcare applications. For instance, Kadime Göğebakan et al. (2024)[28] have developed drug prescription recommendation systems for patients with diabetes and chronic kidney disease, highlighting the adaptability of these technologies to address complex medical needs.

Moreover, federated approaches such as the Personal Health Train, explored by Ananya Choudhury et al. (2024)[20], utilize Semantic Web principles to calculate quality indicators while ensuring patient privacy. Such initiatives illustrate the sophisticated integration of Semantic Web technologies into healthcare practices, enabling a more nuanced analysis of patient data while adhering to strict privacy standards.

Finally, using knowledge graphs to capture complex relationships within EHRs by Bader Aldughayfiq et al. (2023)[10], demonstrates the transformative potential of these technologies. They improve the functionality of EHR systems and enhance the comprehensive analysis of patient outcomes and risk factors.

Together, these studies exemplify the significant advancements in data reuse and knowledge management facilitated by Semantic Web technologies, underscoring their indispensable role in evolving healthcare data systems for better quality care and informed clinical decision-making.

## 5.3 What are the existing knowledge gaps regarding handling health data through Semantic Web technologies?

To answer research question Q3, all selected papers were read thoroughly to find out what method/approach/technology is proposed and their pros and cons. The key points of this analysis are summarized in Tables 7 and 8, (additionally see Table 9).

Our review examined various healthcare informatics methodologies, grouping them by similar technological strategies and analyzing their disadvantages.

#### 5.3.1 *Semantic Web and Ontology Engineering.*

- Limited scalability and adaptability to diverse data types.
- High complexity in integrating and maintaining updated ontologies.
- Dependence on extensive domain knowledge and continuous collaboration with domain experts.

#### 5.3.2 *Data Integration and Security.*

- Privacy concerns when handling sensitive patient data, especially in cloud-based systems.
- Incomplete tools for semantic harmonization, leading to potential data integrity issues.

#### 5.3.3 *Advanced-Data Processing Techniques.*

- Challenges in model explainability, particularly in machine learning applications.
- Limited testing and validation in real-world settings may affect the generalizability of the models.

# 6 DISCUSSION

## 6.1 Opportunities for future research

Our review answered initial research questions by analyzing various healthcare informatics methodologies and their limitations. It highlighted the critical role of semantic web technologies in improving health data management and identified gaps that suggest future research directions. These include developing dynamic methodologies and robust security measures to enhance clinical decision-making and patient care:

#### 6.1.1 *Semantic Web and Ontology Engineering.*

- Developing dynamic ontologies that can adapt to rapidly evolving medical standards without significant manual intervention.
- Automation in mapping local terms to standardized vocabularies to improve data sharing.
- The authors acknowledge that there are not yet many studies connecting semantic web tools with large language models (LLMs), likely due to the newness and fast development of the field. However, they see this as a major research opportunity. Combining LLMs with semantic web technologies has the potential to significantly improve data processing, interoperability, and decision-making in healthcare.

#### 6.1.2 *Data Integration and Security.*

- Development of more robust, transparent, and user-friendly security mechanisms that can be easily audited.
- Efficient real-time data processing techniques to handle the increasing volume of health data without compromising privacy or performance.

#### 6.1.3 *Advanced Data Processing Techniques.*

- Enhancing the transparency and explainability of complex models used in healthcare.
- Creating methods that are readily validated and transferable to diverse clinical settings.

Such an analysis also draws attention to the necessity of further developing information technology in healthcare, primarily through enhanced interoperability and security measures while improving generic data-processing capabilities relevant to health workers.

## 6.2 Limitations of the Scoping Review

This scoping review, while comprehensive in its approach, encountered some limitations that must be acknowledged. Firstly, the inclusion and exclusion criteria may have narrowed the scope of reviewed Semantic Web technologies and their applicability in healthcare. Specifically, the criteria limited the selection to articles explicitly referencing the Semantic Web in their titles or abstracts. This restriction might have led to the exclusion of relevant studies where the term was not explicitly mentioned or was obscured.

Moreover, the review did not incorporate a quality appraisal of the included studies, potentially affecting the overall quality of the evidence presented. The absence of quality assessment could result in the inclusion of lower-quality studies, which might influence the conclusions drawn from the review.

To address these issues and enhance literature coverage, the reference lists of all selected articles were meticulously examined to identify additional significant studies that might have been overlooked. Future research efforts could expand the scope of this review by incorporating studies from other fields and performing comparative analyses to broaden the understanding and applicability of Semantic Web technologies across various domains.

# 7 CONCLUSIONS

The meticulous scoping review of Semantic Web Technologies in healthcare has exhibited that they are used mostly to enhance EHRs. Our findings further show that healthcare systems have improved data interoperability and management using RDF, OWL, and SPARQL technologies. However, it still remains difficult to adopt these technologies widely due to complexities related to technology deployment and integration, which then calls for continued research into their application.

The continuous expansion of Semantic Web usage in the healthcare sector demonstrates how it can radically change medical data management. For instance, linking records and inference capabilities are a good foundation for more advanced health data analysis platforms. In the future, there's a need to make integration procedures easier and improve these technologies' ability to accommodate increasing amounts of information from different healthcare providers.

The current Semantic Web implementations have notable gaps in knowledge, which this review has looked into, especially regarding

**Table 7: Methodology Analysis: Advantages and Disadvantages**

| Ref | Proposed Method/Approach/Technology | Advantages | Disadvantages |
|---|---|---|---|
| [50] | The methodology utilizes a "weighted state transition logic" implemented via semantic web technologies to model clinical pathway changes. This includes defining specific "From" and "To" states, a "transition state" for transitions, and associating weights like duration and cost to transitions. The system generates pathways, detects conflicts, and validates pathways based on these weighted transitions. | The methodology predicts future patient states through modeling state changes and transitions, manages clinical pathways adaptively by anticipating state transitions over time, and integrates duration, cost, comfort, and belief as evaluative measures in transitions, employing semantic web technologies for robust healthcare data integration. | The methodology reacts to changes in the patient's state without proactively predicting future conflicts. Machine learning application within this system is limited by a lack of explainability and constraints on utilizing predictions for adaptive management. The state transition logic is restricted to asserting new states without retracting old ones and generates only a single pathway to the target state, lacking alternatives. |
| [26] | The methodology involves defining clinically relevant variables and data structures using standardized terminologies and EN/ISO 13606 archetypes to identify and normalize EHR data. It includes the creation of multi-layered ontologies for modeling and mapping data to different standards like EN/ISO 13606 and OMOP CDM. The process culminates in inserting normalized EHR extracts into an ontology-based repository and extracting data via SPARQL queries to produce OMOP CDM-compliant tables. | The methodology offers flexibility and adaptability by adding conceptual layers and mapping locally defined concepts to various standards with minimal resource usage. It supports automated data extraction and facilitates the reuse of clinical information through ontological representations, enabling the effortless application of the methodology across different use cases. Its standard-agnostic approach permits transformations between various standards like EN/ISO 13606, OpenEHR, FHIR, OMOP CDM, and i2b2 without the need for database modifications. | The development of the ontology-based clinical repository over many years may not suit institutions needing rapid implementation. Additionally, the tool for inserting data from standardized extracts into ontologies is incomplete, delaying the creation of OMOP CDM tables. |
| [40] | The methodology integrates, anonymizes, and verifies primary and secondary data using Apache Kafka while applying machine learning techniques like KNN for data cleaning and outlier detection. Data is standardized by classifying reliability, translating into SNOMED concepts, mapping to FHIR elements, and transforming into the Holistic Health Record format. | Integrates Semantic Web and machine learning to process data into the Holistic Health Record format enhancing data utilization from EHRs and IoMT devices. Supports the generation of actionable insights and personalized prevention plans, aiding healthcare providers in decision-making. | The main drawback of the methodology is its limited evaluation of a specific pancreatic cancer use case at a single hospital without broader testing across various use cases and data sources. |
| [37] | The study implemented the ABCDE medical guideline using Prova and SHACL rule-based approaches incorporating SPARQL queries to interact with FHIR-RDF data. The guidelines were structured by setting conditions to manage rule execution and the effectiveness was evaluated using 1000 generated FHIR-bundles. | The Prova-based approach offers rule chaining and failure handling with connectivity to multiple SPARQL endpoints for enhanced error reporting and query efficiency. It is convertible to RuleML for broader application reuse. Conversely, the SHACL-based approach excels in performance, requiring less integration, making it ideal for specific scenarios. | Not mentioned (the paper does not explicitly discuss any drawbacks of the proposed methodology approach or technology) |
| [25] | The Oral Health and Disease Ontology (OHD) was collaboratively created by a multidisciplinary team that first defined the domain and necessary terms. Relevant terms were imported from existing ontologies, such as BFO and OGMS, supplemented by new terms when necessary. Additionally, a relation was established to sequentially link patient encounters enhancing the clarity and efficiency of data queries. | The Oral Health and Disease Ontology enhances data understanding through clear terminology and supports multiple coding systems for flexibility. It facilitates structured queries and allows for broad application across various medical domains. Additionally, it supports automated reasoning to derive insights from the data. | The main limitations of the methodology include the need for deep knowledge of vendor-specific database schemas, potential communication barriers with EHR vendors, and the challenge of encouraging the adoption of ontology-based systems over existing vendor-specific frameworks. |
| [56] | The methodology involves creating a network-based framework using FHIR and RDF to represent cancer data, integrating genetic and phenotypical data from EHRs for 1011 patients. Utilizing RDF to generate a patient-genetic-phenotypic network and applying the Node2vec algorithm to extract features. Evaluating cancer prediction using various machine learning models and a feature bagging approach from the FHIR model. | The methodology enhances cancer prediction by accurately predicting cancer using Electronic Health Record data with satisfactory precision. Significantly improving prediction outcomes by integrating genetic information, which supports early diagnosis. Utilizing a network-based framework with FHIR and RDF standards to streamline the cancer prediction process achieving superior performance compared to existing methods for most cancer types. | The limitations of the methodology include challenges in distinguishing between germline and somatic mutations in the genetic data leading to potential biases. The restricted availability of genetic information at some medical facilities limits the adaptability of the best-performing models. A constrained dataset of cancer cases with unknown primaries potentially impacts the robustness of the analysis. |
| [53] | The methodology includes a multi-layer system with user authentication and attribute-based access control utilizing a revocable searchable encryption scheme to secure patient data. Integration of a knowledge graph for storing user and patient data coupled with techniques to perform secure efficient searches on encrypted data. Cloud computing and edge computing principles are used to handle data processing efficiently, reducing the operational load on medical organizations. | The methodology offers consistent data retrieval performance as datasets expand. Capability to manage schema evolution and changes effectively. Enhanced scalability and cost-efficiency through cloud-based computation delegation, including reduced client-side processing needs. | The summary highlights several challenges cloud computing introduces privacy and security risks for healthcare data. Cloud-based EHR systems struggle with performance consistency amid data growth and diverse data types. Previous systems, like those using SWRL by Walid et al., faced scalability and support limitations. The knowledge graphs used lacked essential data and properties for a robust healthcare system. |
| [7] | The methodology involved the ontology development 101 methods focusing on specifying, designing, formalizing, and evaluating the ontology based on over 62000 normalized patient records from a Malaysian rehabilitation center. | The methodology enhances communication between exercise game developers and medical practitioners by aligning exercise game data with SNOMED-CT, facilitates the electronic exchange of clinical data, integrates common rehabilitation assessments, and is developed from real-world data on the MIRA platform, grounding it in practical application. | The methodology's limitation is that it does not incorporate the exercise game rehabilitation ontology with a Clinical Decision Support System (CDSS) crucial for effective patient assessment and clinical decision-making. |

**Table 8: Methodology Analysis: Advantages and Disadvantages - cont**

| Ref | Proposed Method/Approach/Technology | Advantages | Disadvantages |
|---|---|---|---|
| [26] | The methodology involved using RDF Data Cube Vocabulary and SDMX standards to model cancer registry data transforming tabular data into RDF triples and integrating it into the Neo4j graph database. The performance of Neo4j was then compared to a traditional relational database for querying the data. | The approach leveraged existing ontologies to harmonize data across different registries, utilized graph databases for faster queries and more flexible data retrieval compared to relational databases, and aimed to enhance interoperability by potentially releasing data as linked open data. | Not mentioned (the paper does not discuss any drawbacks of the proposed approach) |
| [46] | The study utilized a triplestore to integrate multiple heterogeneous data sources, including EHR claims and various community data, converting them into RDF format. It then linked EHR and claims data deterministically and executed SPARQL queries to identify new chronic opioid users. | The methodology supports consistent data definitions and classifications using ontologies, handles heterogeneous data through its schemaless nature, and efficiently models and queries complex hierarchical clinical data. | The methodology highlights challenges such as the FHIR format generating an excessive number of triples, the impracticality of converting entire clinical data warehouses to RDF, the efficiency of targeting specific data for analysis, and the risk of data loss during transformation, emphasizing the need for careful validation against original sources. |
| [13] | The methodology involved a thorough assessment of patient needs for a personal health library (PHL), identification of necessary technologies, and description of infrastructures like the Solid platform and personal knowledge graphs for managing and integrating knowledge. An initial prototype and its practical application were also developed, aligning the mobile health app's features with identified user requirements. | The approach enhances healthcare by empowering patients and caregivers through control over their data, equipping providers with tools for managing health knowledge, and fostering third-party app development using the PHL. It extends healthcare interaction beyond clinical settings facilitating continuous patient-provider connections and informing new treatment strategies. | Not mentioned (the paper does not discuss any drawbacks of the proposed methodology approach or technology) |
| [24] | The methodology employs a two-part framework utilizing edge computing principles consisting of an internal Access Handler for authentication and decision-making and a Document Processor for handling and encrypting documents. The second part manages data in an untrusted cloud server. | The methodology enhances security using an encrypted access control mechanism suitable for multi-authority environments and employs MA-ABE encryption to manage decryption keys efficiently. It alleviates load bottlenecks and offers a flexible, semantically rich access control system. | The system's limitations include only allowing brute force attribute revocation, lacking patient-end delegation for revoking access, and not supporting keyword searches over encrypted EHR data, highlighting areas for potential enhancements. |
| [36] | The methodology employed an NLQ2Program approach for the MIMIC-SPARQL* EHR-QA dataset defining a custom grammar for exploring the knowledge graph. It generated pseudo-gold programs semi-supervisedly trained a sequence-to-sequence model for translating natural language questions and used ensemble-based uncertainty decomposition to detect ambiguous questions. | The NLQ2Program approach enhances the handling of complex inference tasks and multi-modal data, exceeding typical query language limitations. It achieves performance comparable to advanced NLQ2Query models without complete training data and introduces an ensemble-based uncertainty decomposition to measure question ambiguity, effectively identifying ambiguous input in QA research for the first time. | The proposed methodology lacks a user interface for interactive clarifications and program modifications. Additionally, the data uncertainty metric intended to indicate question ambiguity may inaccurately reflect model errors instead. |
| [28] | The methodology included developing the DIAKID system by sourcing a dataset expert knowledge and clinical guidelines, organizing patient data into a populated DIAKID ontology with new and modified ontologies, and crafting SWRL rules to tailor drug dosage recommendations and warnings for drug interactions and potassium-raising drugs for individual patients. | The approach involves developing a unique DIAKID ontology integrating modified DMTO with new ontologies for drug interactions and patient profiles utilizing SWRL rules for automated drug dose adjustments and interaction warnings and uniquely addressing treatment recommendations for patients with both Type 2 Diabetes Mellitus and Chronic Kidney Disease, including alerts for potassium-increasing drugs. | The proposed system, while innovative, is currently limited to Type 2 Diabetes Mellitus and Chronic Kidney Disease, necessitating expansion for broader applicability to conditions like heart failure and hypertension. It primarily adjusts drug dosages based on kidney function, potentially overlooking other critical factors in different diseases. Furthermore, the system requires enhancements to effectively manage the complexities of drug interactions and dosage adjustments across a more extensive range of chronic conditions. |
| [48] | The study utilized a two-level ontology structure grounded in the OMOP CDM and clinical guidelines to construct a knowledge graph. EHR data was converted into OMOP CDM format and subsequently into RDF triples establishing semantic relationships and a comprehensive patient information model. The system engaged in semantic reasoning on this model to analyze previously unused clinical information, necessitating collaboration from both medical and domain experts throughout its development. | The methodology leverages deductive reasoning and medical evidence to clarify clinical decision-support outcomes, enhancing interpretability for clinicians. It standardizes medical concepts and data within an ontological structure with clear semantic relationships facilitating traceability and acceptance. Furthermore, it constructs a detailed clinical trajectory for patients by analyzing treatment sequences and includes a visualization module to communicate key clinical insights and reasoning processes effectively. | The knowledge graph system is limited to structured EHR data and does not incorporate omics or medical imaging data, which could provide more comprehensive data coverage. Additionally, it faces challenges due to the fragmentation of patient clinical information across multiple hospitals, which complicates the use of multi-center EHR data due to security and cross-institutional reasoning issues. |
| [20] | The methodology involves extracting data from hospital EHR systems, cleaning and converting it to the FAIR data format, storing this data in hospital-based repositories and utilizing the Personal Health Train (PHT) infrastructure to calculate quality indicators while maintaining data privacy by keeping the data within the hospital environment. | The approach effectively reduces errors in EHR data through FAIRification, alleviates the burden of data registration and assembly, and addresses privacy concerns by localizing data and analysis using the Personal Health Train (PHT) infrastructure. This method also facilitates timely calculations of quality indicators and supports ongoing comparative effectiveness research. | The methodology faces several challenges including unforeseen ethical legal and social implications during implementation a lack of stakeholder involvement and complex processes for gaining approvals. Additionally, it struggles with manual data integration due to extraction limitations and raises security and trust concerns among hospital IT departments regarding external computations on local datasets. |
| [9] | The methodology focuses on creating an ontological framework to represent and share clinical pathways standardizing them using SNOMED CT and HL7 standards and developing a unique coding system for digitizing and encoding clinical pathway data. | The methodology enhances clinical pathway management systems by enabling them to operate independently yet in sync with other health systems, integrates data analytics for decision-making, and supports international standardization to improve healthcare quality, reduce costs, and enhance patient outcomes. | The methodology highlights limitations in clinical pathways (CPs) noting they are often seen merely as supports for electronic medical records rather than as central elements of healthcare systems typically developed in non-standardized formats with low adherence to medical terminologies and lacking a dedicated coding system. |

**Table 9: Methodology Analysis: Advantages and Disadvantages - cont**

| Ref | Proposed Method/Approach/Technology | Advantages | Disadvantages |
|---|---|---|---|
| [10] | The methodology involves using Protégé to develop an ontology that defines entities and relationships within the EHR domain processing the MIMIC III dataset with Ontotext Refine to convert data into RDF and represent it graphically and querying the resulting knowledge graph with SPARQL to derive insights. | The approach enhances EHR data integration and analysis by standardizing data representation, improving the efficiency and effectiveness of data analysis for clinical decisions, and contributing to evidence-based knowledge graph development. This advances the field by addressing scalability, interoperability, and clinical validity in knowledge graph development, ultimately aiming to improve patient outcomes and reduce healthcare costs. | The approach faces challenges, including scalability, interoperability, clinical validity, privacy and security risks, limited clinical adoption, and a need for more comprehensive evaluation. |
| [8] | The methodology involves converting various EHR data sources into local OWL ontologies, merging these into a single global ontology, and evaluating the system across diverse data formats. | The methodology unifies heterogeneous EHR data into a single model, enhances data accessibility and error reduction through semantic querying, and improves interoperability by integrating various data formats into a centralized ontological system, potentially leading to better health outcomes and cost reductions. | The methodology faces limitations, including its inability to process unstructured EHR data, lack of a user-friendly graphical interface, unmeasured sensitivity, and difficulties managing uncertain, incomplete, and vague medical knowledge. |
| [55] | The methodology includes semi-automated collection of real-world data from EHRs and genetic reports extraction and normalization of genetic disease and drug data using UMLS integration using the Genetic Testing Ontology and transformation of relational database data into RDF format using the D2RQ tool. | The methodology facilitates the examination of associations between genetic data and treatment decisions, supports precision oncology decision-making, and serves as a pilot for further clinical applications. It utilizes RDF to mathematically model data relationships enabling detailed analysis and advanced graph mining to uncover patterns and insights for drug repurposing. | The methodology faces limitations as it only identifies associative relationships between drugs, diseases, and genes without confirming causality and lacks temporal data in RDF representations, which may introduce biases and hinder the detection of dynamic changes over time. |
| [48] | The methodology involves creating the FHIR-Ontop-OMOP system to generate virtual clinical knowledge graphs from OMOP databases using Ontop technology. It automates the mapping between OMOP CDM and FHIR RDF through a two-step process with Turtle Template Mapping Language and a Java-based converter. The system's effectiveness was tested by assessing data transformation fidelity and ensuring the clinical knowledge graphs conformed to FHIR RDF standards. | The virtual CKG approach efficiently uses existing OMOP databases without needing extra storage and enhances interoperability between FHIR and OMOP CDM standards, demonstrating significant potential for supporting healthcare AI applications. | The mappings between FHIR and OMOP CDM are initial and may need further development particularly in handling vocabulary and encounter type differences. Additionally, the evaluation was restricted to a single OMOP CDM instance with plans for broader testing. |
| [23] | The EMR2vec platform integrates patient data with clinical trials using a vectorization and matching process combining NLP machine learning and semantic web technologies. Key stages include extracting medical terms, transforming EMR and clinical trial data into vectors, reducing dimensionality, and matching patients to trials using similarity measures. The platform notably applies ontological reasoning to align structured and unstructured medical data facilitating the linkage of patients with relevant clinical trials. | The EMR2vec platform integrates EMR data with clinical trials using NLP machine learning and semantic web technologies to accelerate clinical research. This method involves creating a vector space model from a "bag of medical terms" derived from trial criteria, transforming EMR data into vectors, and matching these vectors to trials using dimensionality reduction and orthogonal projection. It provides a comprehensive approach to systematically linking patients with appropriate trials leveraging structured and unstructured medical data. | The proposed methodology has limitations in fully utilizing unstructured clinical text and handling missing EMR fields crucial for matching clinical trial criteria. It lacks mechanisms to standardize and extract laboratory test details essential for effectively linking EMRs to clinical trials impacting its performance in clinical research applications. |
| [44] | The study developed and compared table-based and graph-based EHR QA datasets finding that graph-based approaches using the TREQS model significantly enhanced accuracy by up to 34% over table-based datasets. TREQS incorporated advanced techniques like pointer generation and dynamic attention to improve the translation of natural language questions into SQL, effectively addressing the complexities of querying Electronic Health Records. | The proposed graph-based approach for EHR QA offers significant benefits: it more naturally represents relationships between entities, simplifying queries with SPARQL, which aligns closer to natural language than SQL. Additionally, the TREQS model shows enhanced performance on this dataset, improving relation prediction by 5.1% and accuracy by 3.6%. | The main limitation of the graph-based EHR QA approach is its scalability concerning inference time. As the underlying knowledge graph expands, the processing time for SPARQL queries increases significantly, becoming a bottleneck and resulting in slower response times compared to SQL queries on larger datasets. |
| [52] | The methodology consists of two main components: an Authentication Module that manages access control using user attributes and organizational policies and a Data Processing Module responsible for encryption decryption search token generation, encrypted index creation, and attribute revocation. The system uniformly employs a single revocable searchable attribute-based encryption scheme and handles attribute revocation by updating ciphertext and secondary secret keys within the Data Processing Module. | The system enhances digital health security by integrating advanced security features for cloud EHR systems streamlining the attribute revocation process with updates only needed at the cloud service provider (CSP) level. It also improves the efficiency of keyword searches reducing network latency and client-side computing demands. Additionally, the use of a single encryption scheme simplifies operations, enhancing user-friendliness. | The previous system's drawbacks include the complexity of managing two encryption schemes, slow search times for large datasets, and the absence of an attribute revocation feature, which is critical for adapting to changes in user attributes and organizational policies. |
| [35] | The methodology involves selecting entities for the ontology related to medical and telehealth data for chronic conditions, defining relationships and restrictions to ensure data validity and coherence through expert reviews, and implementing the ontology in OWL-DL. | The proposed methodology provides flexibility and reusability through its ontology-based framework, allowing data model modifications without altering API code. It includes a dynamic REST API that automatically generates endpoints from the ontology, improving adaptability for different applications. Additionally, it integrates linked data and semantic annotations, enhancing data integration and minimizing ambiguity. | The main drawbacks include the immaturity of HL7 FHIR, which necessitates frequent updates to the ontology to reflect changes in FHIR resources, and the placement of constraint checking using SPIN above the persistence layer, which may affect the efficiency of operations. |

dynamic ontology management and real-time data processing. To fill these gaps, we need directed research aimed at creating methods that will adjust according to new medical standards and handle large amounts of information quickly without jeopardizing either its security or performance. On top of that, further investigations into automatic mappings and sophisticated protection systems would also be important for reducing the complications involved.

Using Semantic web technologies in healthcare has given hope for better Electronic Health Record (EHR) systems. One of the serious problems to be solved is how to fully integrate these technologies into current health systems without making radical modifications. Continued innovation and research in this area are necessary if we want to exploit the full benefits of Semantic web technologies, which may result in significant transformations in health informatics and patient care management.

## ACKNOWLEDGMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) Brazil – Finance Code 001 and by the São Paulo Research Foundation (FAPESP grants #2019/07665-4 and #2013/07375-0 CEPID).

## REFERENCES


[1] [n. d.]. About SPARQL - EU Data Portal. https://data.europa.eu/pt/about/sparql. Accessed: 2023-07-10.

[2] [n. d.]. Electronic Health Record. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/electronic-health-record. Accessed: 2023-07-10.

[3] [n. d.]. RDF 1.1 Concepts and Abstract Syntax. https://www.w3.org/TR/rdf12-concepts/. Accessed: 2023-07-10.

[4] [n. d.]. Resource Description Framework (RDF). https://www.w3.org/RDF/. Accessed: 2023-07-10.

[5] [n. d.]. Shapes Constraint Language (SHACL). https://www.w3.org/TR/shacl/. Accessed on 10 July 2024.

[6] [n. d.]. What is RDF? https://www.ontotext.com/knowledgehub/fundamentals/what-is-rdf/. Accessed: 2023-07-10.

[7] Zul Hilmi Abdullah, Lailatul Qadri Zakaria, Waidah Ismail, and Shaharudin Ismail. 2022. Developing an Exercise Games Ontology Towards Standard Rehabilitation Clinical Terms. In *2022 Euro-Asia Conference on Frontiers of Computer Science and Information Technology (FCSIT)*. IEEE, 173–176.

[8] Ebtsam Adel, Shaker El-Sappagh, Sherif Barakat, Kyung Sup Kwak, and Mohammed Elmogy. 2022. Semantic architecture for interoperability in distributed healthcare systems. *IEEE Access* 10 (2022), 126161–126179.

[9] Ayman Alahmar, Matteo Ermando Crupi, and Rachid Benlamri. 2020. Ontological framework for standardizing and digitizing clinical pathways in healthcare information systems. *Computer Methods and Programs in Biomedicine* 196 (2020), 105559.

[10] Bader Aldughayfiq, Farzeen Ashfaq, NZ Jhanjhi, and Mamoona Humayun. 2023. Capturing semantic relationships in electronic health records using knowledge graphs: An implementation using mimic iii dataset and graphdb. In *Healthcare*, Vol. 11. MDPI, 1762.

[11] Edward P Ambinder. 2005. Electronic health records. *Journal of oncology practice* 1, 2 (2005), 57.

[12] Nariman Ammar, James E. Bailey, Robert Lowell Davis, and Arash Shaban-Nejad. 2020. The Personal Health Library: A Single Point of Secure Access to Patient Digital Health Information. *Studies in health technology and informatics* 270 (2020), 448–452. https://api.semanticscholar.org/CorpusID:219991175

[13] Nariman Ammar, James E Bailey, Robert L Davis, Arash Shaban-Nejad, et al. 2021. Using a personal health library–enabled mHealth recommender system for self-management of diabetes among underserved populations: use case for knowledge graphs and linked data. *JMIR formative research* 5, 3 (2021), e24738.

[14] Hilary Arksey and Lisa O'Malley. 2005. Scoping studies: towards a methodological framework. *International journal of social research methodology* 8, 1 (2005), 19–32.

[15] Dörthe Arndt, Ben De Meester, Pieter Bonte, Jeroen Schaballie, Jabran Bhatti, Wim Dereuddre, Ruben Verborgh, Femke Ongenae, Filip De Turck, Rik Van de Walle, et al. 2016. Improving OWL RL reasoning in N3 by using specialized rules. In *Ontology Engineering: 12th International Experiences and Directions Workshop on OWL, OWLED 2015, co-located with ISWC 2015, Bethlehem, PA, USA, October 9-10, 2015, Revised Selected Papers 12*. Springer, 93–104.

[16] Caio Viktor S Avila, Wellington Franco, Amanda DP Venceslau, Tulio Vidal Rolim, Vania MP Vidal, and Valéria M Pequeno. 2021. MediBot: an ontology-based chatbot to retrieve drug information and compare its prices. *Journal of Information and Data Management* 12, 2 (2021).

[17] Tim Berners-Lee, James Hendler, and Ora Lassila. 2001. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific american* 284, 5 (2001), 34–43.

[18] Shinead Casey, Pat Doody, and Andrew Shields. 2022. An ontology-based system for cancer registry data. In *2022 33rd Irish Signals and Systems Conference (ISSC)*. IEEE, 1–6.

[19] Kei-Hoi Cheung, Eric Prud'hommeaux, Yimin Wang, and Susie Stephens. 2009. Semantic Web for Health Care and Life Sciences: a review of the state of the art. *Briefings in bioinformatics* 10, 2 (2009), 111–113.

[20] Ananya Choudhury, Esther Janssen, Bart C Bongers, Nico LU van Meeteren, Andre Dekker, and Johan van Soest. 2024. Colorectal cancer health and care quality indicators in a federated setting using the Personal Health Train. *BMC Medical Informatics and Decision Making* 24, 1 (2024), 121.

[21] Vinícius Costa Lima, Domingos Alves, Filipe Andrade Bernardi, and Rui Pedro Charters Lopes Rijo. 2023. Security approaches for electronic health data handling through the Semantic Web: A scoping review. *Semantic Web* 14, 4 (2023), 771–784.

[22] Ben De Meester, Dörthe Arndt, Pieter Bonte, Jabran Bhatti, Wim Dereuddre, Ruben Verborgh, Femke Ongenae, Filip De Turck, Erik Mannens, and Rik Van de Walle. 2015. Event-driven rule-based reasoning using EYE. In *ISWC2015*. CEUR.

[23] Houssein Dhayne, Rima Kilany, Rafiqul Haque, and Yehia Taher. 2021. EMR2vec: Bridging the gap between patient data and clinical trial. *Computers & Industrial Engineering* 156 (2021), 107236.

[24] Sharad Dixit, Karuna Pande Joshi, Seung Goel Choi, and Lavanya Elluri. 2022. Semantically rich access control in cloud ehr systems based on ma-abe. In *2022 IEEE 8th intl conference on big data security on cloud (BigDataSecurity), IEEE intl conference on high performance and smart computing,(HPSC) and IEEE intl conference on intelligent data and security (IDS)*. IEEE, 1–10.

[25] William D Duncan, Thankam Thyvalikakath, Melissa Haendel, Carlo Torniai, Pedro Hernandez, Mei Song, Amit Acharya, Daniel J Caplan, Titus Schleyer, and Alan Ruttenberg. 2020. Structuring, reuse and analysis of electronic dental data using the Oral Health and Disease Ontology. *Journal of biomedical semantics* 11 (2020), 1–19.

[26] Santiago Frid, Xavier Pastor Duran, Guillem Bracons Cucó, Miguel Pedrera-Jiménez, Pablo Serrano-Balazote, Adolfo Muñoz Carrero, Raimundo Lozano-Rubí, et al. 2023. An ontology-based Approach for consolidating patient data standardized with European Norm/International Organization for standardization 13606 (EN/ISO 13606) into Joint Observational Medical Outcomes Partnership (OMOP) repositories: description of a methodology. *JMIR Medical Informatics* 11, 1 (2023), e44547.

[27] Christian Fürber and Martin Hepp. 2010. Using SPARQL and SPIN for data quality management on the semantic web. In *Business Information Systems: 13th International Conference, BIS 2010, Berlin, Germany, May 3-5, 2010. Proceedings 13*. Springer, 35–46.

[28] Kadime Göğebakan, Ramazan Ulu, Rahib Abiyev, and Melike Şah. 2024. A drug prescription recommendation system based on novel DIAKID ontology and extensive semantic rules. *Health Information Science and Systems* 12, 1 (2024), 27.

[29] Andy Götz. 2023. The fair principles: Trusting in fair data repositories. *Open Access Government* (2023). https://api.semanticscholar.org/CorpusID:259958866

[30] Rafat Hammad, Malek Barhoush, and Bilal H Abed-Alguni. 2020. A Semantic-Based Approach for Managing Healthcare Big Data: A Survey. *Journal of Healthcare Engineering* 2020, 1 (2020), 8865808.

[31] AKM Bahalul Haque, BM Arifuzzaman, Sayed Abu Noman Siddik, Abul Kalam, Tabassum Sadia Shahjahan, TS Saleena, Morshed Alam, Md Rabiul Islam, Foyez Ahmmed, and Md Jamal Hossain. 2022. Semantic web in healthcare: a systematic literature review of application, research gap, and future research avenues. *International Journal of Clinical Practice* 2022, 1 (2022), 6807484.

[32] Pascal Hitzler. 2021. A review of the semantic web field. *Commun. ACM* 64, 2 (2021), 76–83.

[33] Ian Horrocks, Peter F Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosof, Mike Dean, et al. 2004. SWRL: A semantic web rule language combining OWL and RuleML. *W3C Member submission* 21, 79 (2004), 1–31.

[34] Mahtab Karami. 2018. Semantic web: a context for medical knowledge discovering and sharing. *Frontiers in Health Informatics* 7 (2018), 2.

[35] Vassilis Kilintzis, Ioanna Chouvarda, Nikolaos Beredimas, Pantelis Natsiavas, and Nicos Maglaveras. 2019. Supporting integrated care with a flexible data management framework built upon Linked Data, HL7 FHIR and ontologies. *Journal of biomedical informatics* 94 (2019), 103179.

[36] Daeyoung Kim, Seongsu Bae, Seungho Kim, and Edward Choi. 2022. Uncertainty-aware text-to-program for question answering on structured electronic health records. In *Conference on Health, Inference, and Learning*. PMLR, 138–151.

[37] Gerhard Kober, Livio Robaldo, and Adrian Paschke. 2022. Modeling medical guidelines by Prova and SHACL accessing FHIR/RDF. Use case: the medical ABCDE approach. In *dHealth 2022*. IOS Press, 59–66.

[38] Alex Kozlenkov, Rafael Penaloza, Vivek Nigam, Loic Royer, Gihan Dawelbait, and Michael Schroeder. 2006. Prova: Rule-based java scripting for distributed web applications: A case study in bioinformatics. In *Current Trends in Database Technology–EDBT 2006: EDBT 2006 Workshops PhD, DataX, IIDB, IIHA, ICSNW, QLQP, PIM, PaRMA, and Reactivity on the Web, Munich, Germany, March 26-31, 2006, Revised Selected Papers 10*. Springer, 899–908.

[39] Craig S Lockwood and Andrea C. Tricco. 2020. Preparing scoping reviews for publication using methodological guides and reporting standards. *Nursing & health sciences* 22 1 (2020), 1–4. https://api.semanticscholar.org/CorpusID:211725104

[40] George Manias, Ainhoa Azqueta-Alzúaz, Athanasios Dalianis, Jacob Griffiths, Maritini Kalogerini, Konstantina Kostopoulou, Eleftheria Kouremenou, Pavlos Kranas, Sofoklis Kyriazakos, Danae Lekka, et al. 2024. Advanced Data Processing of Pancreatic Cancer Data Integrating Ontologies and Machine Learning Techniques to Create Holistic Health Records. *Sensors* 24, 6 (2024), 1739.

[41] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, and t PRISMA Group*. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine* 151, 4 (2009), 264–269.

[42] Zachary Munn, Micah DJ Peters, Cindy Stern, Catalin Tufanaru, Alexa McArthur, and Edoardo Aromataris. 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC medical research methodology* 18 (2018), 1–7.

[43] Senthil Kumar Narayanasamy, Kathiravan Srinivasan, Yuh-Chung Hu, Satish Kumar Masilamani, and Kuo-Yi Huang. 2022. A contemporary review on utilizing semantic web technologies in healthcare, virtual communities, and ontology-based information processing systems. *Electronics* 11, 3 (2022), 453.

[44] Junwoo Park, Youngwoo Cho, Haneol Lee, Jaegul Choo, and Edward Choi. 2021. Knowledge graph-based question answering with electronic health records. In *Machine Learning for Healthcare Conference*. PMLR, 36–53.

[45] Micah DJ Peters, Christina M Godfrey, Hanan Khalil, Patricia McInerney, Deborah Parker, and Cassia Baldini Soares. 2015. Guidance for conducting systematic scoping reviews. *JBI Evidence Implementation* 13, 3 (2015), 141–146.

[46] Emily R Pfaff, Robert Bradford, Marshall Clark, James P Balhoff, Rujin Wang, John S Preisser, Kellie Walters, and Matthew E Nielsen. 2021. Cohort Identification Using Semantic Web Technologies: Triplestores as Engines for Complex Computable Phenotyping. *medRxiv* (2021), 2021–12.

[47] Thanyalak Rattanasawad, Kanda Runapongsa Saikaew, Marut Buranarach, and Thepchai Supnithi. 2013. A review and comparison of rule languages and rule-based inference engines for the Semantic Web. In *2013 International Computer Science and Engineering Conference (ICSEC)*. IEEE, 1–6.

[48] Yong Shang, Yu Tian, Min Zhou, Tianshu Zhou, Kewei Lyu, Zhixiao Wang, Ran Xin, Tingbo Liang, Shiqiang Zhu, and Jingsong Li. 2021. EHR-oriented knowledge graph system: toward efficient utilization of non-used information buried in routine clinical practice. *IEEE Journal of Biomedical and Health Informatics* 25, 7 (2021), 2463–2475.

[49] Duncan Smith and Tracey Bowden. 2017. Using the ABCDE approach to assess the deteriorating patient. *Nursing standard (Royal College of Nursing (Great Britain) : 1987)* 32 14 (2017), 51–63. https://api.semanticscholar.org/CorpusID:39952415

[50] Hong Sun, Dörthe Arndt, Jos De Roo, and Erik Mannens. 2021. Predicting future state for adaptive clinical pathway management. *Journal of biomedical informatics* 117 (2021), 103750.

[51] Aman Tyagi and Preetvanti Singh. 2019. Health Information System. *Healthcare Policy and Reform* (2019). https://api.semanticscholar.org/CorpusID:5313390

[52] Redwan Walid, Karuna P Joshi, and Seung Geol Choi. 2021. Secure cloud ehr with semantic access control, searchable encryption and attribute revocation. In *2021 IEEE international conference on digital health (ICDH)*. IEEE, 38–47.

[53] Redwan Walid, Karuna Pande Joshi, and Seung Geol Choi. 2024. Leveraging semantic context to establish access controls for secure cloud-based electronic health records. *International Journal of Information Management Data Insights* 4, 1 (2024), 100211.

[54] Guohui Xiao, Emily Pfaff, Eric Prud'hommeaux, David Booth, Deepak K Sharma, Nan Huo, Yue Yu, Nansu Zong, Kathryn J Ruddy, Christopher G Chute, et al. 2022. FHIR-Ontop-OMOP: Building clinical knowledge graphs in FHIR RDF with the OMOP Common data Model. *Journal of biomedical informatics* 134 (2022), 104201.

[55] Yiqing Zhao, Anastasios Dimou, Feichen Shen, Nansu Zong, Jaime I Davila, Hongfang Liu, and Chen Wang. 2022. PO2RDF: representation of real-world data for precision oncology using resource description framework. *BMC medical genomics* 15, 1 (2022), 167.

[56] Nansu Zong, Victoria Ngo, Daniel J Stone, Andrew Wen, Yiqing Zhao, Yue Yu, Sijia Liu, Ming Huang, Chen Wang, Guoqian Jiang, et al. 2021. Leveraging genetic reports and electronic health records for the prediction of primary cancers: algorithm development and validation study. *JMIR Medical Informatics* 9, 5 (2021), e23586.

[57] Lei Zou. 2018. *SPARQL*. Springer New York, New York, NY, 3554–3558. https://doi.org/10.1007/978-1-4614-8265-9_80803