Article

Capturing Semantic Relationships in Electronic Health Records Using Knowledge Graphs: An Implementation Using MIMIC III Dataset and GraphDB

Bader Aldughayfiq [1], Farzeen Ashfaq [2], N. Z. Jhanjhi [2] and Mamoona Humayun [1,*]

1 Department of Information Systems, College of Computer and Information Sciences, Jouf University, Sakaka 72388, Saudi Arabia; bmaldughayfiq@ju.edu.sa
2 School of Computer Science—SCS, Taylor's University, Subang Jaya 47500, Malaysia; farzeen.ashfaq@sd.taylors.edu.my (F.A.); noorzaman.jhanjhi@taylors.edu.my (N.Z.J.)
* Correspondence: mahumayun@ju.edu.sa

Abstract: Electronic health records (EHRs) are an increasingly important source of information for healthcare professionals and researchers. However, EHRs are often fragmented, unstructured, and difficult to analyze due to the heterogeneity of the data sources and the sheer volume of information. Knowledge graphs have emerged as a powerful tool for capturing and representing complex relationships within large datasets. In this study, we explore the use of knowledge graphs to capture and represent complex relationships within EHRs. Specifically, we address the following research question: Can a knowledge graph created using the MIMIC III dataset and GraphDB effectively capture semantic relationships within EHRs and enable more efficient and accurate data analysis? We map the MIMIC III dataset to an ontology using text refinement and Protege; then, we create a knowledge graph using GraphDB and use SPARQL queries to retrieve and analyze information from the graph. Our results demonstrate that knowledge graphs can effectively capture semantic relationships within EHRs, enabling more efficient and accurate data analysis. We provide examples of how our implementation can be used to analyze patient outcomes and identify potential risk factors. Our results demonstrate that knowledge graphs are an effective tool for capturing semantic relationships within EHRs, enabling a more efficient and accurate data analysis. Our implementation provides valuable insights into patient outcomes and potential risk factors, contributing to the growing body of literature on the use of knowledge graphs in healthcare. In particular, our study highlights the potential of knowledge graphs to support decision-making and improve patient outcomes by enabling a more comprehensive and holistic analysis of EHR data. Overall, our research contributes to a better understanding of the value of knowledge graphs in healthcare and lays the foundation for further research in this area.

Keywords: electronic health records; knowledge graphs; semantic relationships; data analysis; MIMIC III; GraphDB; ontology

1. Introduction

Healthcare practitioners and researchers increasingly rely on electronic health records (EHRs), which are digital versions of conventional paper-based medical data. A variety of patient data are contained in EHRs, including demographics, medical history, diagnosis, treatments, and results. In accordance with a study by [1], EHRs enhance patient safety by giving healthcare providers rapid access to patient data, regardless of location or time, which enables them to make more knowledgeable decisions regarding patient care. The fact that EHRs give healthcare workers rapid access to patient data is another significant advantage of using them. Ref. [2] claims that EHRs reduce pharmaceutical errors by giving medical practitioners decision-support tools including drug interaction alerts and dose suggestions. As they may be password-protected and encrypted, EHRs also offer a more

secure access to patient data than conventional paper-based medical records, which lowers the risk of theft or unauthorized access. EHRs also have the advantage of being easily shared across healthcare professionals, which is crucial for patients who visit several experts or receive treatment in several healthcare facilities. EHRs enhance care coordination and lower the risk of medical mistakes. EHRs enhance communication between healthcare professionals and patients, which results in better patient outcomes [3]. In order to create more effective treatment regimens, healthcare providers can use EHRs to find patterns and trends in patient data; hence, enhancing diagnosis by giving them access to patient data from many sources [4–7]. This enables medical professionals to establish diagnoses and create treatment strategies that are more effective. Therefore, from a research perspective, EHRs are a useful data source for clinical trials, epidemiological studies, and healthcare studies. Thanks to them, researchers now have quick and easy access to enormous volumes of patient data, which can be used to identify risk factors, track the development of diseases, and evaluate the effectiveness of different therapies with the potential to improve healthcare quality while reducing costs [8]. Figure 1 depicts a typical patient data flow via an EHR, from data entry through analysis.
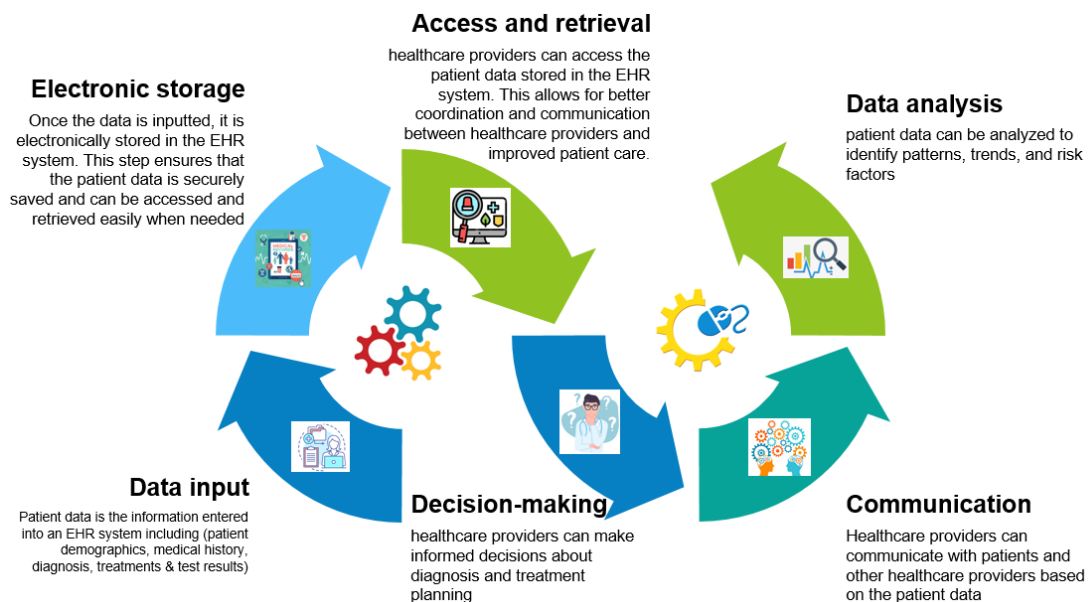

Figure 1. A typical flow of patient data through EHRs, from data input to analysis.

Yet, despite the fact that EHRs have risen in significance as a source of data for researchers and healthcare practitioners, evaluating them can be challenging due to the variety of data sources and the vast quantity of information they contain. EHRs are usually disorganized, hard to analyze, and fragmented. This poses a significant challenge for academics trying to extract useful data from EHRs. The diversity of data sources is one of the biggest obstacles to EHR analysis. Hospitals, clinics, and laboratories are just a few of the places wherein EHRs are routinely gathered. As a result, the structure, nomenclature, and data quality of EHR data can vary. This makes it challenging to efficiently combine data from many sources and analyze them. The standardization of data items, language, and data formats is crucial for effective EHR analysis [9]. The vast amount of data in EHRs presents another difficulty for analysis. Millions of records may be present in EHRs, making it challenging to retrieve pertinent data. This problem is made worse by the fact that EHRs frequently lack a uniform format or coding system, which means they are unstructured. Due to this, it may be challenging to locate and extract particular data points from EHRs [10,11].

To overcome the challenges posed by the heterogeneity of data sources and the unstructured nature of EHRs, researchers have developed new tools and techniques for analyzing

EHRs. These techniques include the use of machine learning algorithms, natural language processing (NLP) techniques, and the development of standardized terminologies and data formats, but while machine learning algorithms and NLP techniques have shown promising results in analyzing EHRs, there are limitations to these methods. For instance, the accuracy of these techniques heavily depends on the quality of the data and the complexity of the task at hand. Moreover, the use of standardized terminologies and data formats may not always be feasible due to the heterogeneity of data sources. In contrast, knowledge graphs have emerged as a promising approach to overcome the limitations of traditional EHR analysis methods. By representing medical knowledge in a structured and semantically rich format, knowledge graphs can facilitate more effective data integration, interoperability, and knowledge discovery. Furthermore, knowledge graphs enable the integration of multiple types of data, including EHRs, clinical guidelines, and biomedical literature, to provide a more comprehensive view of patient health.

To address these limitations, this study aims to design an OWL ontology of the MIMIC III dataset and construct RDF mappings using the ontoText Refine tool [12]. The RDF data will be visualized and queried using graphDB [13] and SPARQL, enabling a more efficient and effective analysis of the data. The clinical validity of the ontology and RDF data will be evaluated through expert review and comparison with existing clinical terminologies. Additionally, privacy and security concerns will be addressed through appropriate data handling and storage. The contributions of this study include:

1. Creating a more standardized and interoperable approach for representing and integrating EHR data.
2. Enabling a more efficient and effective analysis of the data, which can help to identify patterns and relationships that are relevant to clinical decision-making and patient care.
3. Contributing to a more evidence-based approach to knowledge graph development that can improve patient outcomes and reduce healthcare costs.
4. Advancing the field of knowledge graphs for EHR data by addressing key research gaps and contributing to a more scalable, interoperable, and clinically valid approach to knowledge graph development.

In the remainder of this paper, Section 2 describes how this study advances the subject and gives a review of the literature on knowledge graphs in healthcare and EHR analysis. The MIMIC III dataset, the ontology created for it using OWL in Protege, the RDF mapping procedure used to convert the data to the ontology, and the building of the knowledge graph using GraphDB are all described in Sections 3 and 4. Next, Section 5 outlines the study's findings, gives visualizations to highlight the semantic linkages found in the EHR data, and provides examples of how the knowledge graph can be utilized to analyze patient outcomes and spot potential risk factors. Section 6 presents the evaluation of our results. Finally, the study's key conclusions are outlined in Section 7 along with its importance for enhancing EHR analysis and patient outcomes.

2. Literature Review

2.1. Potential of Knowledge Graphs in Healthcare

Machine learning algorithms have been increasingly used to analyze EHRs due to their ability to identify patterns and make predictions based on large and complex datasets. One study found that machine learning algorithms were effective at predicting unplanned hospital readmissions, mortality rates, and length of stay for patients based on EHR data [14]. Another study used machine learning algorithms to develop a predictive model for identifying patients at high risk for developing sepsis, a potentially life-threatening condition [15].

NLP techniques are another type of approach which can be used to extract structured data from unstructured EHRs. By analyzing free-text clinical notes, NLP techniques can identify key clinical concepts, such as diagnoses, procedures, and medications, as well as extract structured data from unstructured sources. A study by [16] found that

NLP techniques were effective at identifying medication-related adverse events from unstructured EHR data.

In recent years, the use of knowledge graphs has gained popularity in various domains due to their ability to represent complex data and relationships between entities. An organized data model known as a knowledge graph captures entities, properties, and relationships among them in a meaningful form. Google first proposed the concept of a knowledge graph in 2012 [17], and it has since gained widespread use in a variety of fields. The concept of knowledge graphs has its roots in the Semantic Web, which was introduced by Tim Berners-Lee in 2001 [18]. The Semantic Web's goal was to build a network of data that both people and machines could access. Hence, a knowledge graph was proposed in order to connect concepts and items in a more logical way, providing consumers with more relevant search results. Since then, they have been adopted and applied in various domains such as sales, logistics, healthcare, security and surveillance, and many others.

Knowledge graphs are also a recent advancement in EHR analysis [19]. A knowledge graph is a structured representation of knowledge that captures relationships between entities, such as diseases, medications, and symptoms [20,21]. By integrating EHR data with external knowledge sources, knowledge graphs can be used to identify complex relationships between clinical concepts and facilitate more accurate predictions. A study by [22] demonstrated the feasibility of using knowledge graphs to identify patients at high risk for hospital readmission.

A knowledge graph is a type of graph-based knowledge representation that uses subject–predicate–object triples to organize information. In a knowledge graph, nodes represent entities, which can include people, places, concepts, and more.The relationships between these entities are represented by edges, which connect the nodes in the graph. Each edge is represented as a subject–predicate–object triple, where the subject is an entity, the predicate is an attribute or relationship, and the object is a value or another entity. Triples are commonly formatted using angle brackets to enclose the subject, predicate, and object, like so: *<subject> <predicate> <object>*.

Triple-based representations of information are common in many scenarios, particularly those involving semantic data modeling. For example, in e-commerce, a triple may represent a product, its price, and a customer's purchase history. In social media, a triple may represent a user, their friend list, and the posts they have made. By representing data as triples, it becomes possible to query and reason about the data in a more structured and efficient way.

A good example of a knowledge graph in action can be found in the MIMIC III dataset, a large electronic health record database. In this dataset, patient information can be represented as subject–predicate–object triples, where the subject is a patient, the predicate is a medical condition or event, and the object is a value or another entity. For example, the triple <Patient #12345> <was_admitted_on> <1 January 2016> may represent a patient's admission date. These triples can be mapped to RDF (Resource Description Framework) triples, a standard for representing data on the web, which enables interoperability with other knowledge graphs and datasets. Figure 2 illustrates how RDF triples can be used to represent patient data in the MIMIC III dataset, where the subject is the patient's unique identifier, the predicate is an attribute or event, and the object is a value or another entity.
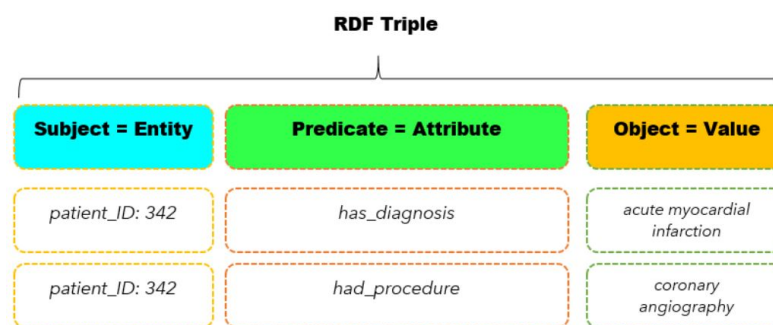
RDF Triple

Subject = Entity	Predicate = Attribute	Object = Value
patient_ID: 342	has_diagnosis	acute myocardial infarction
patient_ID: 342	had_procedure	coronary angiography

Figure 2. The subject–predicate–object format of RDF triples in healthcare data.

Clinicians may easily see the connections between various entities by organizing this data using a knowledge graph, and they can utilize this knowledge to inform their diagnosis and treatment choices. Let us consider the case of a patient who is experiencing chest pain. To diagnose the cause of the chest pain, a physician may need to consider a wide range of possible conditions, such as angina, heart attack, pulmonary embolism, or aortic dissection. Each of these conditions has its own set of symptoms, risk factors, and treatments, and it can be challenging to keep track of all of this information and make a definitive diagnosis. When a patient complains of chest pain and has a history of coronary artery disease, the doctor may swiftly rule out angina as a probable cause of the pain and suggest aspirin therapy as a possible course of treatment. In the discussed example, a node may represent "Angina", which could possibly be connected to other nodes, for example, "Chest Pain", "Shortness of Breath", and "Coronary Artery Disease", as these are all related entities. Similarly, a node representing "Aspirin Therapy" could be connected to nodes representing "Heart Attack", "Stroke", and "Blood Clots", as these are all conditions that can be treated with aspirin.

One of the main advantages of knowledge graphs is that they allow for the integration of external knowledge sources with EHR data. This means that healthcare professionals can draw on a broader range of information to inform their analysis, leading to more accurate and comprehensive insights. Refs. [23,24] found that integrating external knowledge sources with EHR data can help to identify potentially harmful medication combinations that may not be apparent through EHR data alone. Another advantage of knowledge graphs is that they can help to identify previously unknown relationships between clinical concepts. This can be especially valuable for rare or complex conditions wherein traditional data analysis methods may be insufficient. For example, a study by [25] demonstrated the potential of knowledge graphs for identifying new gene–disease associations. However, so far, the research on knowledge graphs for EHR data has been limited by scalability issues, interoperability challenges, clinical validity concerns, and privacy and security risks. Many existing knowledge graph models have been limited to smaller-scale datasets or specific healthcare domains, and have not been widely adopted in clinical practice. Additionally, there is a need for a more rigorous evaluation of knowledge graph models in real-world settings, particularly in terms of their impact on patient outcomes and clinical decision-making. The potential benefits of using knowledge graphs to represent data linked to healthcare are shown in Figure 3.
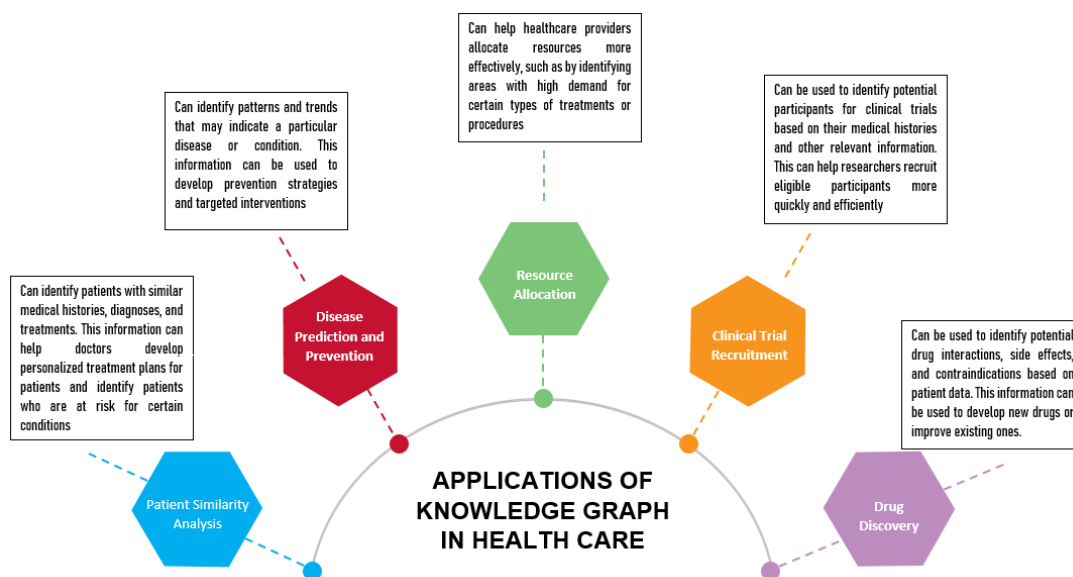
Figure 3. Applications of knowledge graphs in healthcare systems.

In healthcare, knowledge graphs have been used to represent medical knowledge and patient data in a structured way. This has led to the development of clinical decision support systems that provide clinicians with evidence-based recommendations for diagnosis and treatment. Knowledge graphs are also used to represent clinical guidelines, drug interactions, and adverse events, which can help clinicians make informed decisions. They have been increasingly utilized in the healthcare industry to represent and integrate data from various sources. One use case of knowledge graphs in healthcare is clinical decision support systems (CDSSs), which provide physicians with real-time recommendations for diagnosis and treatment based on patient data. The studies conducted by [26,27] demonstrated the effectiveness of using knowledge graphs to develop clinical decision support systems such as for heart failure patients, resulting in a significant improvement in patient outcomes. Another application of knowledge graphs in healthcare is drug discovery [28–30], wherein they are used to integrate and analyze data from various sources, including scientific literature, gene expression, and chemical properties of drugs. A study by [31] used knowledge graphs to identify potential drug targets for Alzheimer's disease, demonstrating the potential for knowledge graphs to accelerate the drug discovery processes.

Furthermore, knowledge graphs have been used for disease surveillance and outbreak prediction. Ref. [32] utilized knowledge graphs to integrate data from various sources, such as social media and public health data, to predict the spread of COVID-19. The knowledge graph provided a unified representation of the data, enabling an accurate prediction of disease outbreaks. Additionally, knowledge graphs have been utilized for patient similarity matching, enabling physicians to identify patients with similar characteristics and medical histories for personalized treatment. A study by [33] developed a knowledge graph-based framework for patient similarity matching, resulting in improved accuracy and efficiency in personalized treatment.

The Electronic Health Record (EHR) system is the central platform wherein all patient data are stored and managed. It contains rich sources of medical data such as patients' medical history, allergies, medications, and diagnoses, which can be represented using knowledge graphs. The use of knowledge graphs in EHRs can help identify relationships between medical concepts and patient outcomes. Furthermore, it can help healthcare providers deliver more personalized and effective care. For example, ref. [34] used a knowledge graph approach to develop a system for identifying patients at risk of readmission, which could help healthcare providers proactively manage patient care. A very similar study was conducted by [35], wherein an approach to enrich EHR data with semantic

annotations to ontologies to build a knowledge graph was developed. The knowledge graph represented a patient's ICU stay in a contextualized manner, which was used by machine learning models to predict 30-days ICU re-admissions. Knowledge graphs can assist in integrating and standardizing data from various sources, including electronic medical records, laboratory results, and medical equipment. Interoperability and data interchange between various healthcare systems and providers can be made easier as a result. For instance, refs. [36,37] employed a knowledge graph approach to combine data from many sources to enhance drug safety monitoring, which could aid healthcare providers in identifying possible drug interactions and negative effects.

With approximately 40,000 patients who were admitted to an intensive care unit, the MIMIC III (Medical Information Mart for Intensive Care III) dataset is a sizable and varied collection of de-identified medical data (ICU). As an illustration, ref. [38] created a knowledge graph to illustrate the connections between therapeutic ideas and patient outcomes in the MIMIC III dataset. Comparing their knowledge graph to conventional machine learning models, they discovered that their knowledge graph significantly increased the accuracy of forecasting patient outcomes. In another study by [39], it was suggested that a graph-based method for electronic health records question answering is more appropriate than a table-based approach. To test their theory, they produced four EHR QA datasets based on a table-based dataset MIMICSQL, and tested a simple Seq2Seq model and a state-of-the-art EHR QA model on all datasets. The graph-based datasets facilitated up to 34% higher accuracy than the table-based dataset without any modification to the model architectures. However, their study had limitations in terms of inference time, especially as the graph size grows and scalability issues as their approach covers a smaller subset of the MIMIC-III dataset.

Ref. [40] outlines a framework for safe medicine recommendations that involves combining patient, disease, and medication information into a single low-dimensional space. The approach converts medication recommendations into a link prediction method while taking into account potential adverse drug responses using a heterogeneous graph made from electronic medical records and medical knowledge graphs. According to experimental findings, SMR provides more accurate recommendations than the state-of-the-art techniques. In another work by [41], the authors propose the use of ontology middleware to integrate IoT healthcare information systems into EHR systems. They argue that the integration process faces challenges due to the lack of interoperability and standardization among different healthcare systems, and propose ontology middleware as a solution to provide a common vocabulary and set of rules for data integration. A study by [42] focuses on curating a domain-specific healthcare knowledge graph for subarachnoid hemorrhage. Another very relevant study is conducted by [43], wherein the authors explore the use of semantic technologies to tackle the interoperability challenges in electronic health records, enabling data integration, reuse, and processing by machine agents as well as propose a transformation of heterogeneous and unstructured patient medical information into a semantic knowledge graph that ensures high levels of interoperability. The pilot study conducted at the UTPL Hospital demonstrated the feasibility of this approach and the potential benefits of structured medical information for doctors, patients, researchers, and governments. However, the authors acknowledge that one of the main challenges to achieving the ambitious objective of managing health data effectively is integrating data from heterogeneous sources and formats. This limitation can be addressed by using large datasets to improve the accuracy and reliability of the semantic models used to represent the medical information. Ref. [44] explains the creation and assessment of a system that produces virtual clinical knowledge graphs (CKGs) from OMOP relational databases. The FHIR–Ontop–OMOP system illustrates the potentials made possible by the compatibility between FHIR and OMOP CDM by exposing the OMOP database as an FHIR-compliant RDF graph. The FHIR Patient, Condition, Procedure, MedicationStatement, Observations, and CodeableConcept instances were present in the CKGs produced from the Medical Information Mart for Intensive Care (MIMIC-III) data source. The paper comes to the

conclusion that CKGs that give a semantic foundation for explainable AI applications in healthcare can be built using the FHIR–Ontop–OMOP architecture. The study's main drawback is that multiple data sources were not reviewed; instead, only one data repository was used.

In conclusion, healthcare organizations continue to have a substantial barrier in integrating and evaluating electronic health record (EHR) data despite the fact that these data are becoming more widely available. A viable method for displaying and integrating EHR data is the use of knowledge graphs, which show the connections between various clinical concepts. However, scale limitations, interoperability difficulties, clinical validity issues, and privacy and security hazards limit the present knowledge graph research employing EHR data. Several of the currently used knowledge graph models are restricted to smaller datasets or particular areas of healthcare and have not found widespread use in clinical settings. More studies concentrated on knowledge graph models in practical contexts are indeed required, especially in terms of their influence on patient outcomes and clinical decision-making.

The aforementioned literature extensively examines the possibilities of utilizing knowledge graphs in the healthcare industry, presenting a range of significant benefits. These benefits encompass improved patient outcomes, enhanced research and development efficacy, and better decision-making regarding diagnosis and treatment. By structuring medical information in a systematic and standardized manner, knowledge graphs can play a vital role in assisting doctors and researchers in comprehending complex data and extracting valuable insights. Ultimately, this comprehensive approach holds the potential to advance patient care and contribute to more effective healthcare practices.

2.2. Use of Knowledge Graphs in Other Domains

Knowledge graphs are being implemented and used more frequently in the sales, marketing, and e-commerce industries through the digital representation of data on all the entities involved (such as items, suppliers, manufacturers, and routes of transportation) and the relationships among them. Knowledge graphs can help these sectors by providing a consistent and organized representation of data, allowing them to better understand their customers, products, and market trends, streamline supply chain management procedures, spot possible bottlenecks, and enhance overall performance [45–51]. For the purpose of assisting consumers in understanding electronic items, ref. [45] presents the idea of a product knowledge graph. The research suggests a sales assistant, which employs semantic advice to aid clients in comprehending the attributes and capabilities of a product. The Internet of Things, for example, can be linked to the product knowledge graph in order to enhance its functionality. Ref. [21] examines the implementation of a semantic content and data value chain for online direct marketing and sales in the travel sector. Two other studies by [52,53] present an ontology and knowledge graph in the area of manufacturing and demand forecasting.

The transportation sector is another business wherein knowledge graphs are increasingly in demand as a means to describe and evaluate complicated data pertaining to traffic patterns, route optimization, and vehicle performance. Ref. [54] created a knowledge graph-based framework for intelligent urban transportation systems. Ref. [55] offers a method based on Semantic Web technology for adhering to EU transport data standards. It transforms information from many Italian and Spanish stakeholders and builds a multi-modal transport knowledge graph for smart querying, exploration, and value-added services.

Table 1 provides a comprehensive summary of various domains wherein knowledge graph applications have been utilized, including healthcare, e-commerce, logistics, sales, marketing, transportation, finance, agriculture, energy, government, and pharmaceuticals, along with the entities, attributes, relationships, ontology, and related literature associated with each domain.

Table 1. Knowledge graph applications in various domains: a comprehensive summary.

Domain	Entities	Attributes	Relationship	Ontology	Related Literature
Healthcare	Patients, Medications, Diseases	Medical History, Drug Effect, Symptoms	Patient Visit Treatment Diagnosis	LOINC SNOMED CT	[56–58] [59–62]
E-commerce	Customers Products Purchasing behavior	Purchases Browsing history Reviews	Retailers Manufacturers Product categories	Schema.org GoodRelations	[48,63,64] [47,49,65]
Logistics	Shipments Warehouses Carriers	Delivery time Cost Performance	Shippers Consignees Shipment locations	W3C ODRL GID	[66–68] [69–71]
Sales	Customers Products Sales channels	Sales volume Revenue Conversion rate	Sales reps distributors Sales regions	Schema.org GoodRelations	[21,45,72] [73–75]
Marketing	Customers Campaigns Channels	Click-through rate Conversion rate Engagement	Advertisers Marketing channels Target demographics	Schema.org FOAF	[21,76] [77–79] [80]
Transportation	Vehicles Routes Traffic patterns	Speed Fuel efficiency Congestion	Transportation modes Geographic locations Traffic flow	SUMO OpenStreetMap ONETT	[81–83] [84–86] [87,88]
Finance	Stocks Investments Market trends	Price Market capitalization Return on investment	Companies Industries Economic indicators	XBRL FIBO	[57,89,90] [91–93]
Agriculture	Crops Soil quality Weather patterns	Yield Quality Nutrient content	Farming practices Weather conditions Soil composition	AgroPortal Agrisemantics Map of Data Standards AgroTagger	[94–96] [53,97]
Energy	Power plants Energy consumption Distribution grids	Energy output Efficiency Emissions	Energy sources Geographic regions Infrastructure	CIM OMS	[98–100] [101–103]
Government	Policies Legislation Public services	Budgets Impact assessments Effectiveness	Government agencies Elected officials Public opinion	Open Government Data FOAF	[104–106] [107–109]
Pharmaceutical	Drugs Diseases Clinical trials	Efficacy Side effects Dosage	Researchers Patients Medical institutions	Drug Ontology NDF-RT	[25,110,111] [112–114]

In conclusion, the use of knowledge graphs in healthcare, particularly in electronic health records (EHRs) and the MIMIC III dataset, has demonstrated their potential in improving patient care and outcomes through the structured representation of medical knowledge and patient data. However, there are still research gaps to be addressed. Existing ontologies for the MIMIC III dataset may have limitations in terms of detail, coverage, and suitability for specific research questions, and the use of semantic technologies may be limited for certain analyses or research questions. Reproducing ontology, mapping, and analysis work on the dataset may also pose challenges, while existing work may be focused on specific research questions or use cases. To address these gaps, future research can focus on creating a more comprehensive and tailored ontology, demonstrating the utility of semantic technologies, providing transparent methodologies for ontology creation, mapping to RDF, and querying using SPARQL, as well as exploring new research questions that have yet to be addressed. Such efforts will further advance our understanding and utilization of knowledge graphs in healthcare and other domains.

3. Methodology

To build an electronic health record (EHR) knowledge graph, we followed a methodology that involved ontology development, data processing, graph representation, and SPARQL querying. Figure 4 illustrates the overall architecture of our methodology, from processing the CSV files to developing the ontology, mapping the data to RDF format, representing it as a graph, and querying it using SPARQL. We first developed an ontology using Protégé to define the entities and relationships in the EHR domain. This allowed us to create a standardized vocabulary that could be used to describe the EHR data. Next, we processed the EHR dataset using Ontotext Refine to map it into RDF format and represent it as a graph. This allowed us to represent the EHR data as a set of nodes and edges, which could be easily queried using SPARQL.

Once the EHR data was represented as a graph, we used SPARQL to query the graph and extract useful information. We developed several use cases (few included in Appendix A) to demonstrate the potential applications of our knowledge graph. Subsequently, we utilized a subset of these use cases to formulate queries that aimed to identify trends and patterns within the data. Through this process, we sought to evaluate the effectiveness and efficiency of our proposed research.
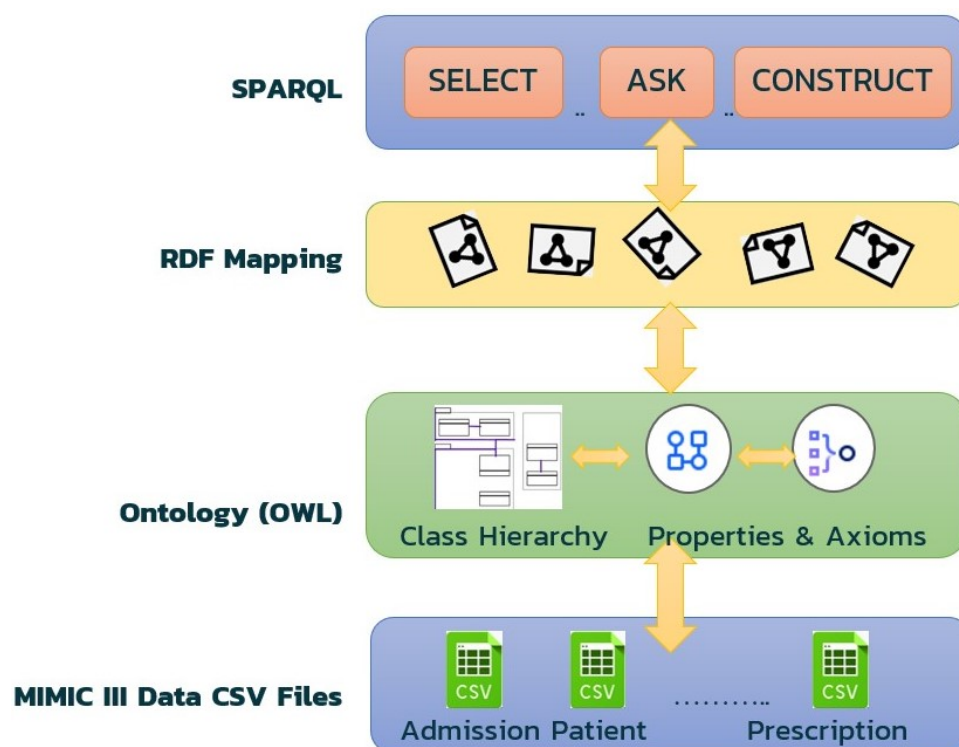

Figure 4. Overall framework of our methodology.

4. Data Selection

For our study, we selected the MIMIC III dataset for building an EHR knowledge graph. The MIMIC III dataset [115] is a large, publicly available database of de-identified electronic medical records from patients admitted to the Beth Israel Deaconess Medical Center between 2001 and 2012. It contains detailed clinical data, including demographic information, diagnoses, procedures, medications, and vital signs, which makes it an ideal dataset for building an EHR knowledge graph. The dataset is provided in the form of CSV (comma-separated values) files, which are organized by clinical domain (e.g., admissions, diagnoses, prescriptions, etc.). There are a total of 26 CSV files in the dataset with sizes ranging from a few kilobytes to several gigabytes. The total size of the dataset is approximately 11 gigabytes.

Before processing the dataset with Ontotext Refine, we first built an ontology using Protégé [116], which defined the entities and relationships between them. We used reasoning to ensure the consistency and completeness of the ontology. Once the ontology was complete, we processed the dataset using Ontotext Refine to map it into RDF format, which allowed us to represent the data as a graph.

4.1. Ontology Development

Ontology is a formal way of representing knowledge that defines a set of concepts and categories, along with their properties and relationships. It provides a standardized vocabulary that can be used to describe the entities and relationships within a domain, such as electronic health records. Ontologies are used in various applications, including semantic search, natural language processing, and knowledge management.

In the ontology development step, we defined the classes, properties, and relationships between the entities in our MIMIC III dataset. We used Protégé, an open-source ontology editor, to create the ontology.

Initially, we established the top-level class, OWLThing *(a built-in class of OWL)*, which forms the basis of the class hierarchy. Then, the major classes such as patient, admission, medication, and diagnosis were defined. We also defined properties that describe the relationships between the entities, such as hasAdmission, hasMedication, and hasDiagnosis. Figure 5 shows the class hierarchy structure we defined on Protégé, where each class and its subclasses are represented. Additionally, Table 2 explains each class and the corresponding files in the MIMIC III dataset. Once the classes have been defined, the next step is to create object properties, which specify the relationships between the classes. In Table 3, instances where properties have two domains, such as "HAS_MEDICATION" with "Admission" and "ICU_Stays", indicate a union of classes as the interpretation. This means that the property can be associated with either the "Admission" class or the "ICU_Stays" class. This allows for flexibility in linking medications to either type of medical encounter. Figure 6 shows the VOWL visualization graph of the object properties, and Table 3 provides information about the domain and range of each object property.

4.2. Demonstration of MIMIC Ontology Instances and Statements

To provide a clearer understanding of the MIMIC ontology, examples of instances and statements using the ontology's classes and properties can be presented. For instance, consider the Patient class, which can have a unique patient ID as an instance. This patient ID can be linked to an Admission class instance using the HAS_ADMISSION property, indicating that the patient has been admitted to the hospital. An Admission instance can further be linked to an ICU Stay instance using the HAS_ICU_STAY property, indicating that the admission involved an ICU stay. The ICU Stay instance can be linked to a Diagnosis instance using the HAS_DIAGNOSIS property, indicating that the ICU stay is associated with a diagnosis. Similarly, the ICU Stay instance can be linked to a Procedure instance using the HAS_PROCEDURE property, indicating that the ICU stay involved a medical procedure.

Moreover, consider the Lab Event class, which can have a specific lab test or measurement as an instance. This lab test or measurement can be linked to a Lab Item class instance using the HAS_LAB_ITEM property, indicating that the lab event is associated with a specific lab item. The Lab Event instance can also be linked to a Patient, Admission, or ICU Stay class instance using the HAS_LAB_RESULT property, indicating that the lab event is associated with a patient, admission, or ICU stay. Similarly, the Medication class can have a specific drug or medical device as an instance, which can be linked to an Individual Item class instance using the HAS_INDIVIDUAL_ITEM property, indicating that the medication involves a specific item.

Figure 5. Class hierarchy of ontology in Protege.

4.3. RDF Mapping

For this research, we used the Ontotext Refine tool to construct RDF triples from the MIMICIII dataset. RDF triples are a way of representing information using three components: subject, predicate, and object. For example, in the context of the MIMICIII dataset, the subject may represent a patient, the predicate may be a characteristic or event related to that patient, and the object is a value or description of that characteristic or event.

After creating the OWL ontology for the MIMIC-III dataset, which defined the classes, properties, and relationships between them and provided a structured vocabulary to represent the data elements, we used the Ontotext Refine tool to generate RDF triples for each CSV file in the dataset. To do this, we first mapped the columns in the CSV file to the properties in the OWL ontology using a mapping file. To make the mapping process easier, we defined a prefix for our RDF subject, predicate, and object, using the shortform "mc". Additionally, we defined a base IRI of "http://mimicIII.com/base/." to ensure that all of our URIs were unique and consistent. Once the mapping file is created, we can use the Ontotext Refine tool to generate RDF triples for each row in the CSV file. The tool uses the mapping file to create triples in the format of subject–predicate–object (SPO), wherein the subject is the unique identifier for each row, the predicate is the property from the OWL ontology, and the object is the value of the data element in the CSV file. To further illustrate the process of RDF mapping, we provide two examples in Figures 7 and 8. The first subfigure shows the RDF mappings for the Admission.csv file, while the second subfigure shows the RDF mappings for the Prescription.csv file.
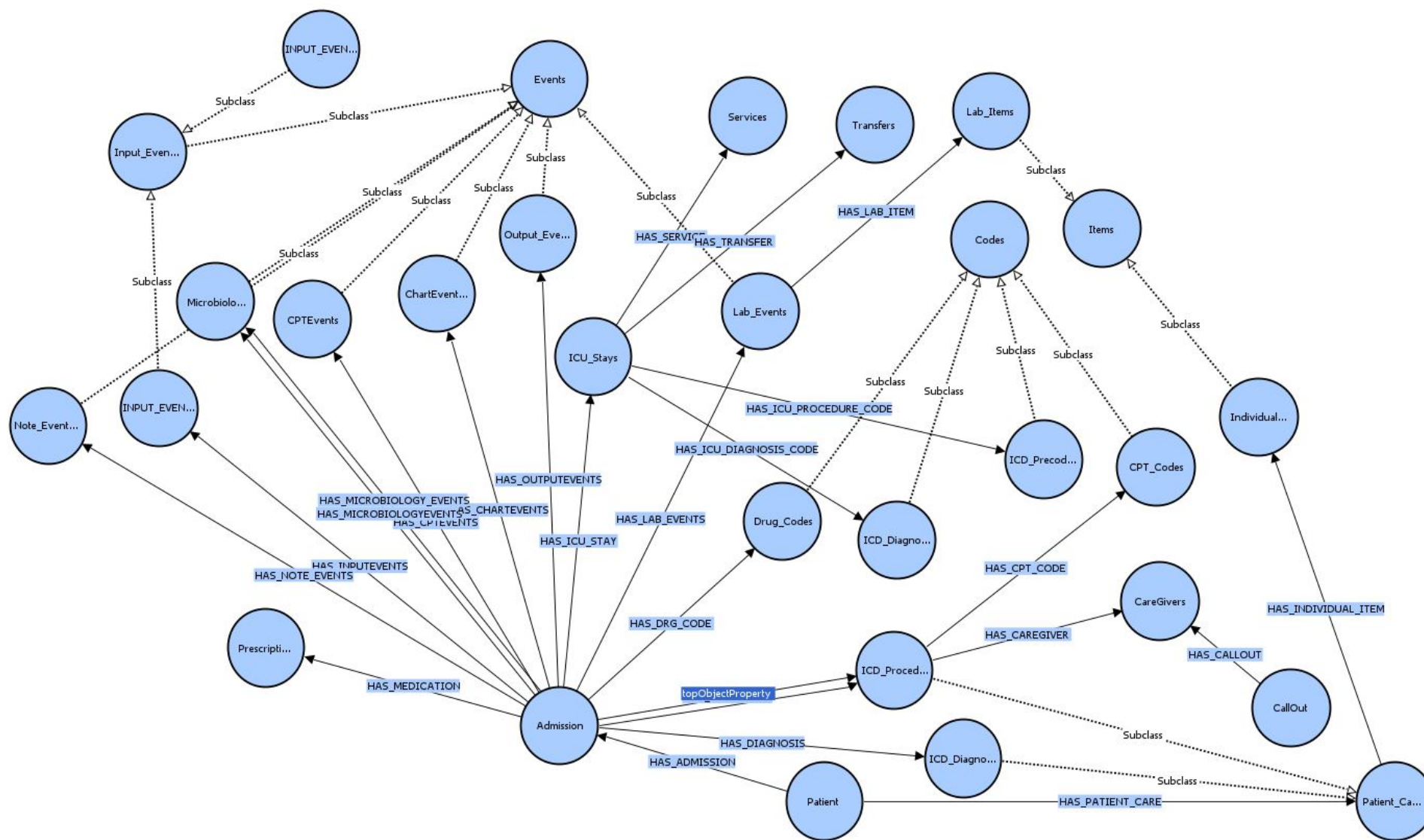
Figure 6. Visualisation of Ontology with object properties using VOWL.

Table 2. Explanation of the class hierarchy as defined in Protege.

Class	Description	Related CSV File
Patient	Information about the patients' demographics, such as age, gender, ethnicity, and marital status is recorded using datatype properties attached to this class	PATIENTS.CSV
Admission	Information about the admission and discharge dates, as well as details about the patient's medical condition and treatment recorded using datatype properties attached to this class	ADMISSIONS.CSV
CareGivers	Information about the caregivers responsible for a patient's care recorded using datatype properties attached to this class	CAREGIVERS.CSV
Patient Care(ICD_Diagnosis, ICD_Procedure	Information about patient's medical surgeries, interventions and medical conditions recorded using datatype properties attached to this class	ICD_Dignosis.CSV, ICD_Procedures.CSV
Codes (CPT, Drug, ICD Dignosis and ICD Procedures)	Information about description of all codes recorded using datatype properties attached to this class	D_CPT.CSV,D_ICD_DIAGNOSES.CSV, D_ICD_PROCEDURES.CSV,DRGCODES.CSV
ICU_Stays	Information about ICU stay, including admission and discharge dates, length of stay, and ICU type recorded using datatype properties attached to this class	ICUSTAYS.CSV
Events (Chart, CPT, Input, Lab, Microbiology, Note, and Output)	Information about all clinical and procedural events and observations recorded using datatype properties attached to the subclass	CHARTEVENTS.CSV, CPTEVENTS.CSV, INPUTEVENTS_CV.CSV, INPUTEVENTS_MV.CSV, LABEVENTS.CSV, NOTEEVENTS.CSV, MICROBIOLOGY.CSV, OUTPUTEVENTS.CSV
Transfers	Information about patient transfers between hospital locations recorded using datatype properties attached to this class	TRANSFERS.CSV
Services	Information about hospital services provided to the patient during their hospital admission recorded using datatype properties attached to this class	SERVICES.CSV
Prescription	Information about medications prescribed to patients recorded using datatype properties attached to this class	PRESCRIPTION.CSV
Callout	Information about patient requests for consultations recorded using datatype properties attached to this class	CALLOUT.CSV

Table 3. Object properties and domains in EHR ontology.

Property	Domain	Range	
HAS_ADMISSION	Patient	Admission	admissions.csv
HAS_ICU_STAY	Admission	ICU_Stays	icustays.csv
HAS_DIAGNOSIS	Admission, ICU_Stays	Diagnosis	diagnoses_icd.csv
HAS_PROCEDURE	Admission, ICU_Stays	Procedure	procedures_icd.csv
HAS_MEDICATION	Admission, ICU_Stays	Prescription	prescriptions.csv
HAS_LAB_EVENTS	Patient, Admission, ICU_Stays	Lab_Events	labevents.csv
HAS_NOTE	Patient, Admission, ICU_Stays	Note_Events	noteevents.csv
HAS_TRANSFER	Admission, ICU_Stays	Transfer	transfers.csv

Table 3. *Cont.*

| Property	Domain	Range	
HAS_SERVICE	Admission, ICU_Stays	Service	services.csv
HAS_LAB_ITEM	Lab_Events	Lab_Items	d_labitems.csv
HAS_INDIVIDUAL_ITEM	Medication, Procedure	Individual Item	d_items.csv
HAS_CAREGIVER	Patient, Admission, ICU_Stay, Procedure	Caregiver	caregivers.csv
HAS_CPT_CODE	Procedure	CPT Code	d_cpt.csv
HAS_DRG_CODE	Admission	DRG Code in the Codes class	drgcodes.csv
HAS_ICU_PROCEDURE_CODE	ICU_Stays	ICU Procedure	d_icd_procedures.csv
HAS_ICU_DIAGNOSIS_CODE	ICU_Stays	ICU Diagnosis Code	d_icd_diagnoses.csv
HAS_PATIENT_CARE	Patient	Patient Care	patient.csv
HAS_ICD_DIAGNOSIS	Patient Care	ICD Diagnosis	diagnoses_icd.csv
HAS_ICD_PROCEDURE	Patient Care	ICD Procedure	procedures_icd.csv
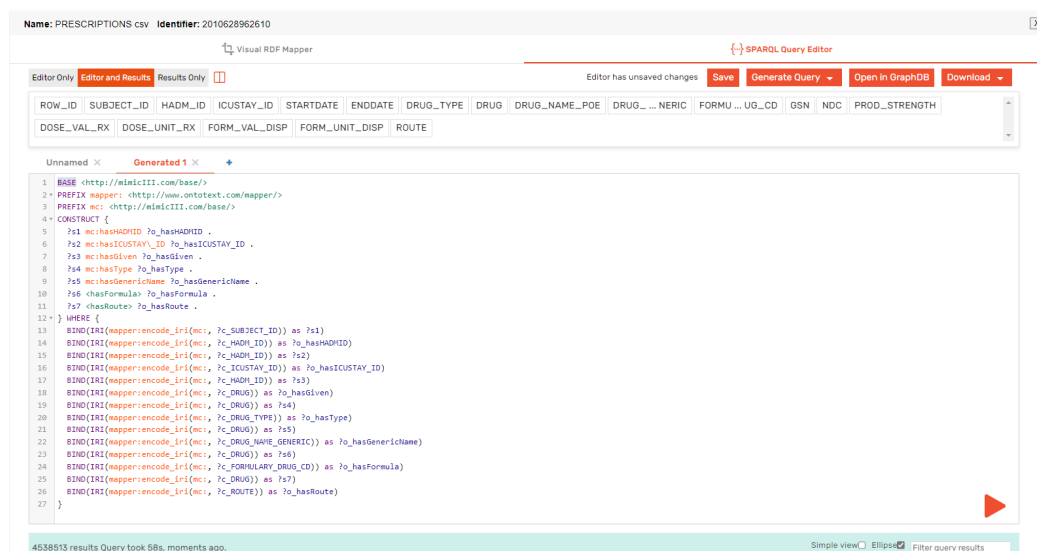
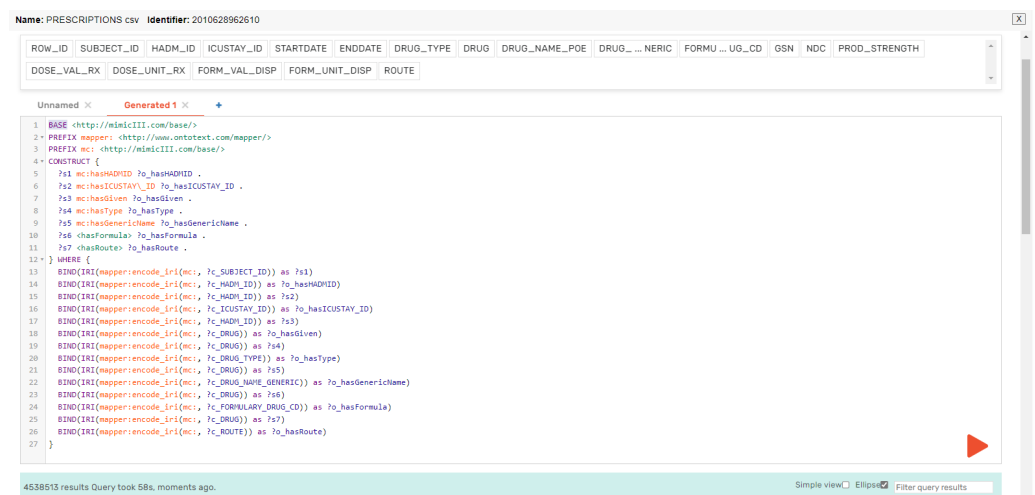**Figure 7.** RDF mapping for Admission.csv.

Figure 8. RDF mapping for Prescription.csv.

To generate the RDF mappings, we used a SPARQL construct query in GraphDB which retrieved the data from the CSV files and mapped it to the OWL ontology. The generated RDF triples were then visualized using the visual RDF mapper in GraphDB. Figures 9 and 10 show an example of the generated RDF graph for the Admission and Prescription tables.

The queries construct statements using the MIMIC ontology's properties and classes. The PREFIX lines define namespace prefixes for the query, which are used to simplify the code and make it easier to read. The BASE line defines the base URI for the MIMIC ontology. The CONSTRUCT block specifies the RDF triples to be created in the new graph. The subject of each triple is a variable that starts with a "?". The predicate and object of each triple are properties and values that belong to the MIMIC ontology. The WHERE block contains a SERVICE clause that specifies the SPARQL endpoint which the data will be retrieved from. The BIND statements assign IRIs to the variables based on the mappings defined in the query. The resulting IRIs are used as the subject, predicate, and object of the RDF triples constructed in the CONSTRUCT block. Finally, queries retrieve data from a SPARQL endpoint and use it to construct new RDF graphs based on the MIMIC ontology.
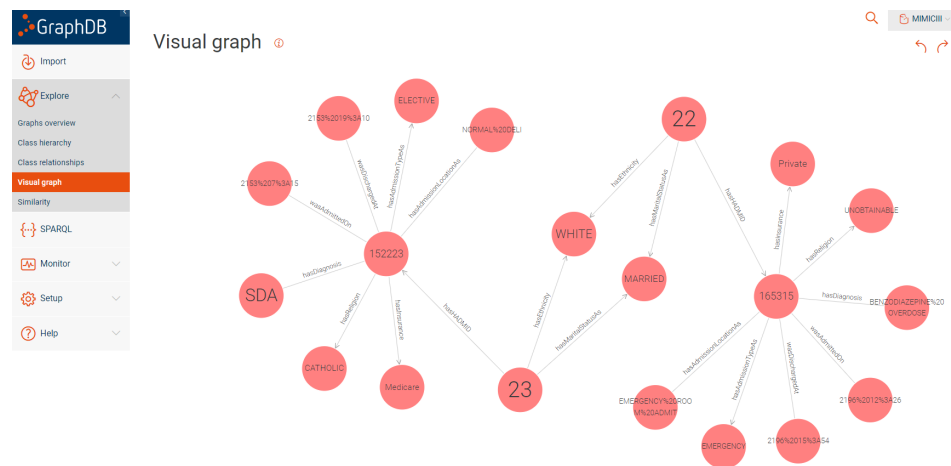

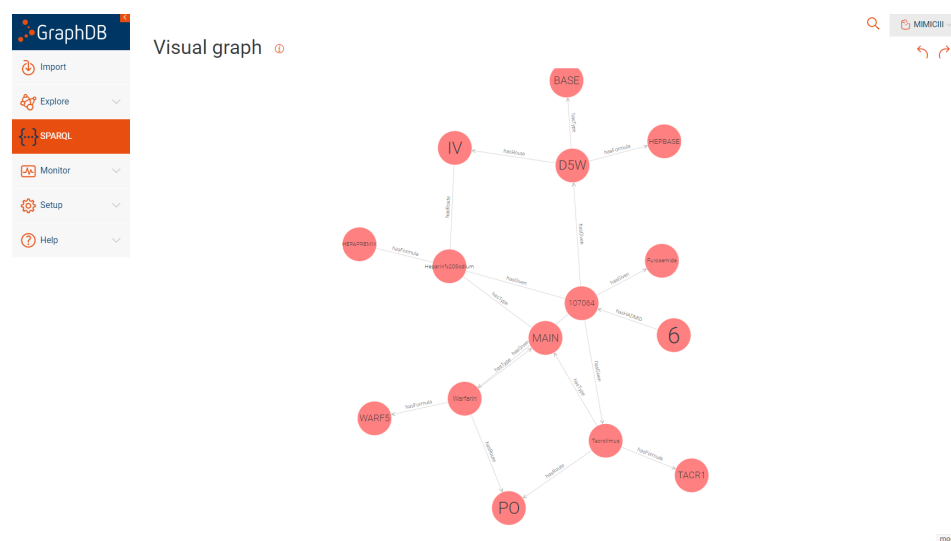**Figure 9.** Sample visual graph for admission.

Figure 10. Sample visual graph for prescription.

This process allowed us to convert the MIMIC-III dataset into a linked data format which can be queried using SPARQL queries and linked to other datasets in the Semantic Web.

5. Results

SPARQL is a powerful query language used for querying RDF data. In this study, we explored the effectiveness of SPARQL queries on the MIMIC-III dataset for extracting meaningful insights. Some of the SPAQRL queries we performed are outlined below.

5.1. Finding Patients with Diabetes

In order to find patients with diabetes in the MIMIC-III dataset, we executed the following SPARQL query as shown in Figure 11:

```
PREFIX mc: <http://mimicIII.com/base/>
PREFIX mapper: <http://www.ontotext.com/mapper/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

SELECT ?patient
WHERE {
  ?patient mc:hasDiabetes "true"^^xsd:boolean .
}
```

Figure 11. Querying for all diabetes patients.

5.2. Finding Patients Who Have Been Diagnosed with Both Hypertension and Diabetes

In this scenario, as shown in Figure 12, we want to find patients who have been diagnosed with both hypertension and diabetes. We have two data files, one containing patient data and another containing diagnosis data. The patient data file includes the patient's ID and date of birth, while the diagnosis data file includes the diagnosis ID, patient ID, and diagnosis code.

```
PREFIX mc: <http://mimicIII.com/base/>
PREFIX mapper: <http://www.ontotext.com/mapper/>

SELECT ?patientID ?dob
WHERE {
  ?patient mc:hasID ?patientID.
  ?patient mc:dateOfBirth ?dob.
  ?diagnosis mc:hasPatientId ?patientID.
  ?diagnosis mc:hasDiagnosisCode "I10".
  ?diagnosis2 mc:hasPatientId ?patientID.
  ?diagnosis2 mc:hasDiagnosisCode "E11".
}
```

Figure 12. Querying for all diabetes and hypertension patients.

5.3. Finding Patients Who Have Been Admitted to the ICU Multiple Times

In this scenario, as shown in Figure 13, we want to find patients who have been admitted to the ICU multiple times. We have two data tables, one containing patient data and another containing ICU stay data. The patient data file includes the patient's ID and date of birth, while the ICU stay data file includes the stay ID, patient ID, and admission date.

```
PREFIX mapper: <http://www.ontotext.com/mapper/>

SELECT ?patientID ?dob
WHERE {
  ?patient mc:hasID ?patientID .
  ?patient mc:dateOfBirth ?dob .
  ?icuStay mc:hasPatientID ?patientID .
  ?icuStay mc:admissionDate ?admissionDate .
  ?icuStay2 mc:hasPatientID ?patientID .
  ?icuStay2 mc:admissionDate ?admissionDate2 .
  FILTER(?admissionDate < ?admissionDate2) .
}
GROUP BY ?patientID
```

Figure 13. Querying for all patients admitted to ICU multiple times.

Overall, our SPARQL queries demonstrate the flexibility and power of semantic technologies for querying large and complex datasets such as MIMIC-III. By leveraging RDF and SPARQL, we were able to easily and effectively search for specific patient populations based on various medical conditions and procedures. These queries can provide valuable insights for clinical research and decision-making.

6. Discussion

6.1. Query Performance Evaluation

In this study, we evaluated the query performance of our proposed ontology-based approach. We used a dataset from the MIMIC-III clinical database and executed a set of sample queries. However, it is important to note that our study is preliminary in nature and our focus was on demonstrating the usefulness and feasibility of the proposed ontology and knowledge graphs in this domain. Therefore, we used a limited set of sample queries to test the performance of the ontology. Specifically, we used queries related to patient demographics, diagnoses, and medications. The average query execution time ranges are less than 0.15 s, which we believe is a significant improvement over existing approaches. However, we acknowledge that query performance may vary depending on the size and complexity of the data, as well as the specific queries used.

6.2. Comparison with Existing Approaches

After constructing the ontology and loading the data into GraphDB, we tested the performance of the system by executing various SPARQL queries. We compared our approach with a traditional relational database management system, MySQL. In MySQL, we had to define queries for each search task, which could become complex and time-consuming for larger datasets. In contrast, with GraphDB, we were able to easily navigate the ontology and data through the graph visualization, which allowed for a more intuitive and user-friendly experience. For example, by clicking on the "Pneumonia" node in the graph, all the patients with pneumonia were immediately displayed without having to run an explicit query. This feature has the potential to greatly enhance the usability of the system in a healthcare setting, where physicians and researchers may not have extensive experience with database query languages. Table 4 presents a average execution time comparison of simple queries from one table executed on GRAPHDB and MySQL.

Table 4. Query execution time comparison between GraphDB and MySQL using MIMIC-III dataset.

Database	Query	Execution Time (s)
GraphDB	SELECT ?patient WHERE { ?patient mc:gender mc:Male . ?patient mc:race mc:White . ?patient mc:marital_status mc:Married }	0.11
MySQL	SELECT * FROM PATIENTS WHERE gender='M' AND race='White' AND marital_status='MARRIED'	1.33
GraphDB	SELECT ?diagnosis WHERE { ?diagnosis mc:icd9_code "41401" }	0.15
MySQL	SELECT * FROM DIAGNOSES_ICD WHERE icd9_code='41401'	1.21
GraphDB	SELECT ?patient ?caregiver WHERE {?patient rdf:type :Patient .?patient :hasCaregiver ?caregiver . ?caregiver :cgid "16175" .}	0.44
MySQL	SELECT p.*FROM Patients p JOIN Caregivers c ON p.CaregiverID = c.CaregiverID WHERE c.CGID = 16175;	1.01

The bar chart in Figure 14 illustrates the stark difference in query execution times between GraphDB and MySQL, with GraphDB consistently demonstrating superior performance. The significantly shorter execution times in GraphDB highlight its efficacy in efficiently processing queries, showcasing its advantage over MySQL in terms of query efficiency.
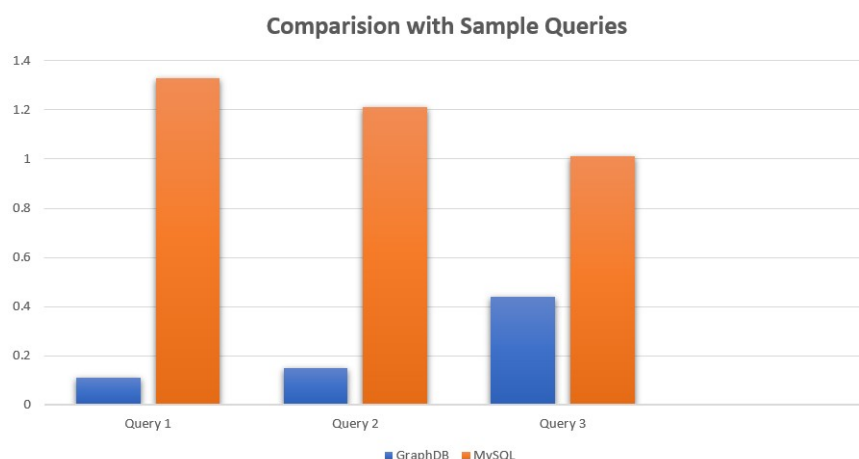
Figure 14. Bar chart illustrating the difference in query execution times between GraphDB and MySQL using three sample queries.

6.3. Interoperability

Interoperability is a key advantage of our proposed approach, as it allows for a seamless integration with other clinical data sources and knowledge bases. Our ontology is based on standard semantic web technologies, such as RDF and OWL, which facilitate data integration and knowledge sharing.

We acknowledge that the interoperability of our approach may depend on the availability and quality of other clinical data sources and ontologies. However, we believe that our approach provides a baseline framework for integrating and harmonizing heterogeneous clinical data sources, which is a critical need in the field.

7. Conclusions

Our work on implementing an EHR knowledge graph using the MIMIC III dataset, GraphDB, and ontology created with Protégé has several significant contributions to the healthcare industry. Firstly, our study demonstrates the immense potential of knowledge graphs in capturing and visualizing complex interactions in EHRs, enabling healthcare practitioners to discover patterns, risk factors, and adverse medication responses. Moreover, our implementation of the EHR knowledge graph significantly reduces the time required to perform queries, with an average query execution time of less than 0.15 s. This improvement in query performance can greatly enhance decision-making in healthcare settings, leading to more efficient and effective patient care.

Furthermore, our study provides a framework for developing automated ontology building techniques, which can significantly reduce the time and effort required to create ontologies for different EHR databases. This development can potentially overcome the significant limitation of subject-matter expertise required for ontology building, which has traditionally limited the scalability of EHR knowledge graphs. By expanding the EHR knowledge graph to include patient-generated data, genetic data, and socioeconomic determinants of health, we can gain a more comprehensive understanding of patient health and provide personalized medication. In terms of limitations, it is worth noting that ensuring interoperability with external datasets and applications is an important consideration for the practical implementation of our EHR knowledge graph. As a next step, we plan to explore linking elements of our ontology and knowledge graph with standard external vocabularies (such as SNOMED CT and LOINC) to achieve this interoperability. This development will allow us to expand the scope of our EHR knowledge graph and potentially enable the integration of data from different EHR databases, ultimately leading to more comprehensive patient care.

Looking ahead, our exploration of machine learning algorithms to detect new risk factors and forecast patient outcomes represents a significant future contribution to the

healthcare industry. This development has the potential to revolutionize decision-making in healthcare settings, enabling healthcare practitioners to identify at-risk patients earlier and provide more personalized and effective treatments. With the continued development and expansion of EHR knowledge graphs, we believe that the potential for improving patient outcomes in the healthcare industry is immense.

In summary, our work on implementing an EHR knowledge graph and demonstrating its potential to capture and visualize complex interactions in EHRs, significantly improve query performance, and develop automated ontology building techniques, has significant contributions to the healthcare industry. The EHR knowledge graph has the potential to revolutionize decision-making in healthcare settings, leading to more efficient and effective patient care, ultimately leading to better patient outcomes.

Author Contributions: Data curation, B.A.; Funding acquisition, B.A.; Investigation, F.A.; Methodology, B.A.; Project administration, M.H.; Supervision, N.Z.J.; Validation, F.A.; Writing—original draft, F.A.; Writing—review editing, N.Z.J. and M.H. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The following table provides a selection of sample use cases that were developed to showcase the potential applications of our knowledge graph. These use cases were designed to demonstrate the capabilities of the graph in capturing and representing complex relationships within Electronic Health Records (EHRs). It is important to note that, while some of these use cases were implemented and executed as RDF tuples to test the functionality of the knowledge graph, they represent a subset of the possible scenarios that can be explored using the graph.

Table A1. Sample use cases demonstrating the potential applications of the knowledge graph in electronic health records (EHRs).

Use Case	Description	Example	Potential Benefits for Physicians
Patient Demographics	Retrieve demographic information for patients, including age, gender, ethnicity, and marital status	Identify all patients with White ethnicity, aged between 30 and 40 years, and married	Better understanding of the patient population they are treating
Length of Stay Analysis	Analyze the duration of hospital stays for patients in different admission categories or with specific diagnoses	Calculate the average length of stay for patients admitted with a diagnosis of sepsis in the Intensive Care Unit (ICU)	Assess the effectiveness of treatment protocols and make data-driven decisions regarding resource allocation and discharge planning

Table A1. *Cont.*

Use Case	Description	Example	Potential Benefits for Physicians
Disease Diagnosis Tracking	Track and analyze the prevalence and distribution of different diagnoses across patient admissions	Determine the frequency of the diagnosis "Acute Myocardial Infarction" across different age groups and genders	Gain insights into the prevalence and distribution of specific diseases within their patient population
Medication Prescriptions	Examine medication prescription patterns and identify commonly prescribed drugs for specific conditions	Identify the most commonly prescribed medication for patients diagnosed with diabetes in the outpatient setting	Utilize this information to ensure adherence to evidence-based treatment

References

1. Tubaishat, A. The effect of electronic health records on patient safety: A qualitative exploratory study. *Inform. Health Soc. Care* **2019**, *44*, 79–91. [CrossRef]
2. Rudolf, J.; Dighe, A. Decision support tools within the electronic health record. *Clin. Lab. Med.* **2019**, *39*, 197–213. [CrossRef]
3. Quinn, M.; Forman, J.; Harrod, M.; Winter, S.; Fowler, K.; Krein, S.; Gupta, A.; Saint, S.; Singh, H.; Chopra, V. Electronic health records, communication, and data sharing: Challenges and opportunities for improving the diagnostic process. *Diagnosis* **2019**, *6*, 241–248. [CrossRef] [PubMed]
4. Chen, J.; Wei, W.; Guo, C.; Tang, L.; Sun, L. Textual analysis and visualization of research trends in data mining for electronic health records. *Health Policy Technol.* **2017**, *6*, 389–400. [CrossRef]
5. Wang, J.; Ouyang, D.; Hom, J.; Chi, J.; Chen, J. Characterizing electronic health record usage patterns of inpatient medicine residents using event log data. *PLoS ONE* **2019**, *14*, e0205379. [CrossRef] [PubMed]
6. Humayun, M.; Jhanjhi, N.Z.; Almotilag, A. Real-time security health and privacy monitoring for Saudi highways using cutting-edge technologies. *Appl. Sci.* **2022**, *12*, 2177. [CrossRef]
7. Almufareh, M.F.; Tehsin, S.; Humayun, M.; Kausar, S. A Transfer Learning Approach for Clinical Detection Support of Monkeypox Skin Lesions. *Diagnostics* **2023**, *13*, 1503. [CrossRef]
8. Rudin, R.; Friedberg, M.; Shekelle, P.; Shah, N.; Bates, D. Getting value from electronic health records: Research needed to improve practice. *Ann. Intern. Med.* **2020**, *172*, S130–S136. [CrossRef]
9. Hersh, W.R.; Weiner, M.G.; Embi, P.J.; Logan, J.R.; Payne, P.R.O.; Bernstam, E.V.; Lehmann, H.P.; Hripcsak, G.; Hartzog, T.H.; Cimino, J.J.; et al. Caveats for the use of operational electronic health record data in comparative effectiveness research. *Med. Care* **2013**, *51*, S30. [CrossRef]
10. Shah, S.; Khan, R. Secondary use of electronic health record: Opportunities and challenges. *IEEE Access* **2020**, *8*, 136947–136965. [CrossRef]
11. Mathew, P.; Pillai, A. Big Data solutions in Healthcare: Problems and perspectives. In Proceedings of the 2015 International Conference on Innovations in Information, Embedded Furthermore, Communication Systems (ICIIECS), Coimbatore, India, 19–20 March 2015; pp. 1–6.
12. Ontotext. OntoText Refine . Available online: https://ontotext.com/products/refine/ (accessed on 10 May 2023).
13. Ontotext. GraphDB. Available online: https://graphdb.ontotext.com/ (accessed on 10 May 2023).
14. Zhou, S.; Lyons, R.; Rahman, M.; Holborow, A.; Brophy, S. Predicting hospital readmission for campylobacteriosis from electronic health records: A machine learning and text mining perspective. *J. Pers. Med.* **2022**, *12*, 86. [CrossRef] [PubMed]
15. Lauritsen, S.; Kalør, M.; Kongsgaard, E.; Lauritsen, K.; Jørgensen, M.; Lange, J.; Thiesson, B. Early detection of sepsis utilizing deep learning on electronic health record event sequences. *Artif. Intell. Med.* **2020**, *104*, 101820. [CrossRef]
16. Edgcomb, J.; Zima, B. Machine learning, natural language processing, and the electronic health record: Innovations in mental health services research. *Psychiatr. Serv.* **2019**, *70*, 346–349. [CrossRef] [PubMed]
17. Blog, G. Introducing the Knowledge Graph: Thing, Not Strings. 2012. Available online: https://blog.google/products/search/introducing-knowledge-graph-things-not/ (accessed on 10 May 2023) .
18. Berners-Lee, T.; Hendler, J.; Lassila, O. The semantic web. *Sci. Am.* **2001**, *284*, 34–43. [CrossRef]
19. Sachdeva, S.; Bhalla, S. Using knowledge graph structures for semantic interoperability in electronic health records data exchanges. *Information* **2022**, *13*, 52. [CrossRef]
20. Hogan, A.; Blomqvist, E.; Cochez, M.; D'Amato, C.; Melo, G.; Gutierrez, C.; Kirrane, S.; Gayo, J.; Navigli, R.; Neumaier, S.; et al. Knowledge graphs. *ACM Comput. Surv.* **2021**, *54*, 1–37 [CrossRef]
21. Fensel, D.; Simsek, U.; Angele, K.; Huaman, E.; Kärle, E.; Panasiuk, O.; Toma, I.; Umbrich, J.; Wahler, A. *Knowledge Graphs*; Springer: Berlin/Heidelberg, Germany, 2020.
22. Wanyan, T.; Honarvar, H.; Azad, A.; Ding, Y.; Glicksberg, B. Deep learning with heterogeneous graph embeddings for mortality prediction from electronic health records. *Data Intell.* **2021**, *3*, 329–339. [CrossRef]
23. Joshi, P.; Masilamani, V.; Mukherjee, A. A knowledge graph embedding based approach to predict the adverse drug reactions using a deep neural network. *J. Biomed. Inform.* **2022**, *132*, 104122. [CrossRef]

24. Zhang, F.; Sun, B.; Diao, X.; Zhao, W.; Shu, T. Prediction of adverse drug reactions based on knowledge graph embedding. *BMC Med. Inform. Decis. Mak.* **2021**, *21*, 1–11. [CrossRef]
25. Nicholson, D.; Greene, C. Constructing knowledge graphs and their biomedical applications. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 1414–1428. [CrossRef]
26. Sheng, M.; Hu, Q.; Zhang, Y.; Xing, C.; Zhang, T. A data-intensive CDSS platform based on knowledge graph. In Proceedings of the Health Information Science: 7th International Conference, HIS 2018, Cairns, Australia, 5–7 October 2018; pp. 146–155.
27. Chaoyu, Z.; Lei, L. A Review of Medical Decision Supports Based on Knowledge Graph. *Data Anal. Knowl. Discov.* **2020**, *4*, 26–32.
28. Zeng, X.; Tu, X.; Liu, Y.; Fu, X.; Su, Y. Toward better drug discovery with knowledge graph. *Curr. Opin. Struct. Biol.* **2022**, *72*, 114–126. [CrossRef] [PubMed]
29. MacLean, F. Knowledge graphs and their applications in drug discovery. *Expert Opin. Drug Discov.* **2021**, *16*, 1057–1069. [CrossRef] [PubMed]
30. Bonner, S.; Barrett, I.; Ye, C.; Swiers, R.; Engkvist, O.; Hoyt, C.; Hamilton, W. Understanding the performance of knowledge graph embeddings in drug discovery. *Artif. Intell. Life Sci.* **2022**, *2*, 100036. [CrossRef]
31. Gao, Z.; Ding, P.; Xu, R. Kg-predict: A knowledge graph computational framework for drug repurposing. *J. Biomed. Inform.* **2022**, *132*, 104133. [CrossRef] [PubMed]
32. Wu, Z.; Xue, R.; Shao, M. Knowledge graph analysis and visualization of AI technology applied in COVID-19. *Environ. Sci. Pollut. Res.* **2022**, *29*, 26396–26408. [CrossRef]
33. Lin, Z.; Yang, D.; Yin, X. Patient similarity via joint embeddings of medical knowledge graph and medical entity descriptions. *IEEE Access* **2020**, *8*, 156663–156676. [CrossRef]
34. Lu, Q.; Nguyen, T.; Dou, D. Predicting patient readmission risk from medical text via knowledge graph enhanced multiview graph convolution. In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual, 14–15 July 2021; pp. 1990–1994.
35. Carvalho, R.; Oliveira, D.; Pesquita, C. Knowledge Graph Embeddings for ICU readmission prediction. *BMC Med. Inform. Decis. Mak.* **2023**, *23*, 12. [CrossRef]
36. Wu, X.; Duan, J.; Pan, Y.; Li, M. Medical knowledge graph: Data sources, construction, reasoning, and applications. *Big Data Min. Anal.* **2023**, *6*, 201–217. [CrossRef]
37. Mythili, R.; Parthiban, N.; Kavitha, V. Construction of heterogeneous medical knowledge graph from electronic health records. *J. Discret. Math. Sci. Cryptogr.* **2022**, *25*, 921–930. [CrossRef]
38. Hou, W.; Zheng, W.; Sheng, M.; Ren, P.; Zuo, B.; Hu, Z.; Liu, X.; Duan, Y. Medical Knowledge Graph Construction Based on Traceable Conversion. In Proceedings of the Health Information Science: 11th International Conference, HIS 2022, Virtual, 28–30 October 2022; pp. 243–257.
39. Park, J.; Cho, Y.; Lee, H.; Choo, J.; Choi, E. Knowledge graph-based question answering with electronic health records. *Mach. Learn. Healthc. Conf.* **2021**, *149*, 36–53.
40. Gong, F.; Wang, M.; Wang, H.; Wang, S.; Liu, M. SMR: Medical knowledge graph embedding for safe medicine recommendation. *Big Data Res.* **2021**, *23*, 100174. [CrossRef]
41. Alqahtani, S.; Alghamdi, M.; Alsaleem, S.; Alshahrani, S. Ontology Middleware for Integration of IoT Healthcare Information Systems in EHR Systems. *Computers* **2018**, *7*, 51.
42. Malik, K.; Krishnamurthy, M.; Alobaidi, M.; Hussain, M.; Alam, F.; Malik, G. Automated domain-specific healthcare knowledge graph curation framework: Subarachnoid hemorrhage as phenotype. *Expert Syst. Appl.* **2020**, *145*, 113120. Available online: https://www.sciencedirect.com/science/article/pii/S0957417419312366 (accessed on 2 March 2023). [CrossRef]
43. Calva, M.; Piedra, N. Interoperability of electronic health records using Semantic Knowledge Graphs. A use case applied at the UTPL University Hospital. *CLEI Electron. J.* **2021**, *24*, 2. [CrossRef]
44. Wu, Y.; Chen, Y.; Wei, W.; Xu, J.; Xu, K.; Zhang, Y.; Xie, G. FHIR-Ontop-OMOP: Building clinical knowledge graphs in FHIR RDF with the OMOP Common data Model. *J. Biomed. Inform.* **2021**, *118*, 103831.
45. Kim, H. Towards a sales assistant using a product knowledge graph. *J. Web Semant.* **2017**, *46*, 14–19. [CrossRef]
46. Jain, S. Exploiting knowledge graphs for facilitating product/service discovery. *arXiv* **2020**, arXiv:2010.05213.
47. Fu, Z.; Xian, Y.; Zhu, Y.; Zhang, Y.; Melo, G. Cookie: A dataset for conversational recommendation over knowledge graphs in e-commerce. *arXiv* **2020**, arXiv:2008.09237.
48. Xu, D.; Ruan, C.; Korpeoglu, E.; Kumar, S.; Achan, K. Product knowledge graph embedding for e-commerce. In Proceedings of the 13th International Conference on Web Search and Data Mining, Houston, TX, USA, 3–7 February 2020; pp. 672–680.
49. Yu, C.; Wang, W.; Liu, X.; Bai, J.; Song, Y.; Li, Z.; Gao, Y.; Cao, T.; Yin, B. FolkScope: Intention Knowledge Graph Construction for Discovering E-commerce Commonsense. *arXiv* **2020**, arXiv:2211.08316.
50. Jain, S. Product discovery utilizing the semantic data model. In *Multimedia Tools and Applications*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 1–27.
51. Li, X.; Lyu, M.; Wang, Z.; Chen, C.; Zheng, P. Exploiting knowledge graphs in industrial products and services: A survey of key aspects, challenges, and future perspectives. *Comput. Ind.* **2021**, *129*, 103449. [CrossRef]
52. Rožanec, J.; Zajec, P.; Kenda, K.; Novalija, I.; Fortuna, B.; Mladenić, D. XAI-KG: Knowledge graph to support XAI and decision-making in manufacturing. In Proceedings of the Advanced Information Systems Engineering Workshops: CAiSE 2021 International Workshops, Melbourne, Australia, 28 June–2 July 2021; pp. 167–172.

53. Yahya, M.; Breslin, J.; Ali, M. Semantic web and knowledge graphs for industry 4.0. *Appl. Sci.* **2021**, *11*, 5110. [CrossRef]
54. Tan, J.; Qiu, Q.; Guo, W.; Li, T. Research on the construction of a knowledge graph and knowledge reasoning model in the field of urban traffic. *Sustainability* **2021**, *13*, 3939. [CrossRef]
55. Scrocca, M.; Comerio, M.; Carenini, A.; Celino, I. Turning transport data to comply with EU standards while enabling a multimodal transport knowledge graph. In Proceedings of the Semantic Web—ISWC 2020: 19th International Semantic Web Conference, Athens, Greece, 2–6 November 2020; pp. 411–429.
56. Zhang, Y.; Sheng, M.; Zhou, R.; Wang, Y.; Han, G.; Zhang, H.; Xing, C.; Dong, J. HKGB: An inclusive, extensible, intelligent, semi-auto-constructed knowledge graph framework for healthcare with clinicians' expertise incorporated. *Inf. Process. Manag.* **2020**, *57*, 102324. [CrossRef]
57. Abu-Salih, B. Domain-specific knowledge graphs: A survey. *J. Netw. Comput. Appl.* **2021**, *185*, 103076. [CrossRef]
58. Li, L.; Wang, P.; Yan, J.; Wang, Y.; Li, S.; Jiang, J.; Sun, Z.; Tang, B.; Chang, T.; Wang, S.; et al. Real-world data medical knowledge graph: Construction and applications. *Artif. Intell. Med.* **2020**, *103*, 101817. [CrossRef] [PubMed]
59. Chen, I.; Agrawal, M.; Horng, S.; Sontag, D. Robustly extracting medical knowledge from EHRs: A case study of learning a health knowledge graph. In *Pacific Symposium on Biocomputing 2020*; World Scientific: Singapore, 2019; pp. 19–30.
60. Tamašauskaitė, G.; Groth, P. Defining a knowledge graph development process through a systematic review. *ACM Trans. Softw. Eng. Methodol.* **2023**, *32*, 1–40. [CrossRef]
61. Celebi, R.; Uyar, H.; Yasar, E.; Gumus, O.; Dikenelli, O.; Dumontier, M. Evaluation of knowledge graph embedding approaches for drug-drug interaction prediction in realistic settings. *BMC Bioinform.* **2019**, *20*, 1–14. [CrossRef]
62. Aldughayfiq, B.; Ashfaq, F.; Jhanjhi, N.; Humayun, M. YOLO-Based Deep Learning Model for Pressure Ulcer Detection and Classification. *Healthcare* **2023**, *11*, 1222. [CrossRef]
63. Zhang, W.; Deng, S.; Chen, M.; Wang, L.; Chen, Q.; Xiong, F.; Liu, X.; Chen, H. Knowledge graph embedding in e-commerce applications: Attentive reasoning, explanations, and transferable rules. In Proceedings of the 10th International Joint Conference on Knowledge Graphs, Virtual, 6–8 December 2021; pp. 71–79.
64. Zhang, W.; Wong, C.; Ye, G.; Wen, B.; Zhang, W.; Chen, H. Billion-scale pre-trained e-commerce product knowledge graph model. In Proceedings of the 2021 IEEE 37th International Conference on Data Engineering (ICDE), Chania, Greece, 19–22 April 2021; pp. 2476–2487.
65. Li, F.; Chen, H.; Xu, G.; Qiu, T.; Ji, F.; Zhang, J.; Chen, H. AliMeKG: Domain knowledge graph construction and application in e-commerce. In Proceedings of the 29th ACM International Conference On Information & Knowledge Management, Virtual, 19–23 October 2020; pp. 2581–2588.
66. Sternberg, E.; Atzmueller, M. Knowledge-based mining of exceptional patterns in logistics data: Approaches and experiences in an Industry 4.0 context. In *Foundations of Intelligent Systems, Proceedings of the 24th International Symposium, ISMIS 2018, Limassol, Cyprus, 29–31 October 2018*; Springer International Publishing: Berlin/Heidelberg, Germany, 2018; pp. 67–77.
67. Zhao, Z.; Zhang, M.; Chen, J.; Qu, T.; Huang, G. Digital twin-enabled dynamic spatial-temporal knowledge graph for production logistics resource allocation. *Comput. Ind. Eng.* **2022**, *171*, 108454. [CrossRef]
68. Garofalo, M.; Pellegrino, M.; Altabba, A.; Cochez, M. Leveraging knowledge graph embedding techniques for industry 4.0 use cases. In *Cyber Defence in Industry 4.0 Systems Furthermore, Related Logistics Furthermore, IT Infrastructures*; IOS Press: Amsterdam, The Netherlands, 2018; pp. 10–26.
69. Bretones Cassoli, B.; Jourdan, N.; Metternich, J. Knowledge Graphs for Data Furthermore, Knowledge Management in Cyber-Physical Production Systems. In Proceedings of the Conference on Production Systems Furthermore, Logistics: CPSL 2022, Vancouver, BC, Canada, 17–20 May 2022; pp. 445–454.
70. Li, J.; Yang, X.; Guan, Y.; Pan, Z. Prediction of Drug–Target Interaction Using Dual-Network Integrated Logistic Matrix Factorization and Knowledge Graph Embedding. *Molecules* **2022**, *27*, 5131. [CrossRef] [PubMed]
71. Listl, F.; Fischer, J.; Beyer, D.; Weyrich, M. Knowledge Representation in Modeling and Simulation: A survey for the production and logistic domain. In Proceedings of the 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria, 8–11 September 2020; Volume 1, pp. 1051–1056.
72. Wang, Z.; Yang, L.; Lei, Z.; Haq, A.; Zhang, D.; Yang, S.; Francis, A. An entity-weights-based convolutional neural network for large-sale complex knowledge embedding. *Pattern Recognit.* **2022**, *131*, 108841. [CrossRef]
73. Ramzy, N.; Ehm, H.; Wiens, V.; Kohnen, L. The Digital Reference: Semantically Connecting Semiconductor Supply Chains to Customers—The Open Online Sales and Marketing Vision. In Proceedings of the CASE Conference, Lyon, France, 23–27 August 2021.
74. Yipa, H.; Liub, Y.; Shetha, A. Using Contact, Content, and Context in Knowledge-Infused Learning: A Case Study of Non-Sequential Sales Processes in Sales Engagement Graphs. In Proceedings of the Knowledge Graph Conference 2021 Workshop on Knowledge-Infused Learning (K-IL), Online, 3 May 2021.
75. Noy, N.; Gao, Y.; Jain, A.; Narayanan, A.; Patterson, A.; Taylor, J. Industry-scale Knowledge Graphs: Lessons and Challenges: Five diverse technology companies show how it's done. *Queue* **2019**, *17*, 48–75. [CrossRef]
76. Dadoun, A.; Troncy, R.; Defoin-Platel, M.; Petitti, R.; Solano, G. Optimizing Email Marketing Campaigns in the Airline Industry using Knowledge Graph Embeddings. *IEEE Data Eng. Bull.* **2021**, *44*, 51–64.
77. Guedea-Noriega, H.; García-Sánchez, F. Integroly: Automatic Knowledge Graph Population from Social Big Data in the Political Marketing Domain. *Appl. Sci.* **2022**, *12*, 8116. [CrossRef]

78. Wang, J.; Chen, B.; Wang, Y.; Xu, Z.; Zhao, W. Research on Intelligent Power Marketing Inspection Model Based on Knowledge Graph. *Sci. Program.* **2022**, *2022*, 7116988. [CrossRef]

79. Sequeda, J.; Briggs, W.; Miranker, D.; Heideman, W. A pay-as-you-go methodology to design and build enterprise knowledge graphs from relational databases. In *The Semantic Web-ISWC 2019, Proceedings of the 18th International Semantic Web Conference, Auckland, New Zealand, 26–30 October 2019*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 526–545.

80. Muzafar, S.; Jhanjhi, N.; Khan, N.; Ashfaq, F. DDoS Attack Detection Approaches in on Software Defined Network. In Proceedings of the 2022 14th International Conference on Mathematics, Actuarial Science, Computer Science Furthermore, Statistics (MACS), Karachi, Pakistan, 12–13 November 2022; pp. 1–5.

81. Ke, P.; Ji, H.; Ran, Y.; Cui, X.; Wang, L.; Song, L.; Zhu, X.; Huang, M. Jointgt: Graph-text joint representation learning for text generation from knowledge graphs. *arXiv* **2021**, arXiv:2106.10502.

82. Pei, S.; Yu, L.; Zhang, X. Improving cross-lingual entity alignment via optimal transport. In Proceedings of the International Joint Conferences on Artificial Intelligence, Macao, China, 10–16 August 2019.

83. Jiang, W.; Luo, J. Graph neural network for traffic forecasting: A survey. *Expert Syst. Appl.* **2022**, *207*, 117921. [CrossRef]

84. Petric Maretic, H.; El Gheche, M.; Chierchia, G.; Frossard, P. GOT: An optimal transport framework for graph comparison. *Adv. Neural Inf. Process. Syst.* **2019**, *32*. [CrossRef]

85. Liu, H.; Li, T.; Hu, R.; Fu, Y.; Gu, J.; Xiong, H. Joint representation learning for multi-modal transportation recommendation. *Proc. AAAI Conf. Artif. Intell.* **2019**, *33*, 1036–1043. [CrossRef]

86. Yuan, H.; Li, G. A survey of traffic prediction: From spatio-temporal data to intelligent transportation. *Data Sci. Eng.* **2021**, *6*, 63–85. [CrossRef]

87. Humayun, M.; Ashfaq, F.; Jhanjhi, N.; Alsadun, M. Traffic management: Multi-scale vehicle detection in varying weather conditions using yolov4 and spatial pyramid pooling network. *Electronics* **2022**, *11*, 2748. [CrossRef]

88. Ashfaq, F.; Ghoniem, R.; Jhanjhi, N.; Khan, N.; Algarni, A. Using Dual Attention BiLSTM to Predict Vehicle Lane Changing Maneuvers on Highway Dataset. *Systems* **2023**, *11*, 196. [CrossRef]

89. Chen, Z.; Wang, Y.; Zhao, B.; Cheng, J.; Zhao, X.; Duan, Z. Knowledge graph completion: A review. *IEEE Access* **2020**, *8*, 192435–192456. [CrossRef]

90. Chen, X.; Jia, S.; Xiang, Y. A review: Knowledge reasoning over knowledge graph. *Expert Syst. Appl.* **2020**, *141*, 112948. [CrossRef]

91. Liu, W.; Zhou, P.; Zhao, Z.; Wang, Z.; Ju, Q.; Deng, H.; Wang, P. K-bert: Enabling language representation with knowledge graph. *Proc. AAAI Conf. Artif. Intell.* **2020**, *34*, 2901–2908. [CrossRef]

92. Long, J.; Chen, Z.; He, W.; Wu, T.; Ren, J. An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in Chinese stock exchange market. *Appl. Soft Comput.* **2020**, *91*, 106205. [CrossRef]

93. Deng, S.; Zhang, N.; Zhang, W.; Chen, J.; Pan, J.; Chen, H. Knowledge-driven stock trend prediction and explanation via temporal convolutional network. In Proceedings of the Companion Proceedings of the 2019 World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; pp. 678–685.

94. Xiaoxue, L.; Xuesong, B.; Longhe, W.; Bingyuan, R.; Shuhan, L.; Lin, L. Review and trend analysis of knowledge graphs for crop pest and diseases. *IEEE Access* **2019**, *7*, 62251–62264. [CrossRef]

95. Liu, Y. DKG-PIPD: A Novel Method About Building Deep Knowledge Graph. *IEEE Access* **2021**, *9*, 137295–137308. [CrossRef]

96. Haussmann, S.; Seneviratne, O.; Chen, Y.; Ne'eman, Y.; Codella, J.; Chen, C.; McGuinness, D.; Zaki, M. FoodKG: A semantics-driven knowledge graph for food recommendation. In *The Semantic Web-ISWC 2019, Proceedings of the 18th International Semantic Web Conference, Auckland, New Zealand, 26–30 October 2019*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 146–162.

97. Drury, B.; Fernandes, R.; Moura, M.; de Andrade Lopes, A. A survey of semantic web technology for agriculture. *Inf. Process. Agric.* **2019**, *6*, 487–501. [CrossRef]

98. Liu, Y.; Wang, G.; Zhou, Y.; Liu, Y. Advanced Technology Evolution Pathways of Nanogenerators: A Novel Framework Based on Multi-Source Data and Knowledge Graph. *Nanomaterials* **2022**, *12*, 838. [CrossRef]

99. Sun, R.; Cao, X.; Zhao, Y.; Wan, J.; Zhou, K.; Zhang, F.; Wang, Z.; Zheng, K. Multi-modal knowledge graphs for recommender systems. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management, Virtual, 19–23 October 2020; pp. 1405–1414.

100. Chen, G.; Hou, J.; Liu, C.; Hu, K.; Wang, J. Visualization Analysis of Cross Research between Big Data and Construction Industry Based on Knowledge Graph. *Buildings* **2022**, *12*, 1812. [CrossRef]

101. Ji, S.; Pan, S.; Cambria, E.; Marttinen, P.; Philip, S. A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 494–514. [CrossRef]

102. Sun, Z.; Huang, J.; Hu, W.; Chen, M.; Guo, L.; Qu, Y. Transedge: Translating relation-contextualized embeddings for knowledge graphs. In *The Semantic Web–ISWC 2019, Proceedings of the 18th International Semantic Web Conference, Auckland, New Zealand, 26–30 October 2019*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 612–629.

103. Li, Q.; Dai, R.; Wei, S.; Zhang, J.; Deng, S. Data Security Knowledge Graph for Active Distribution Network. In *Artificial Intelligence Furthermore, Robotics, Proceedings of the 7th International Symposium, ISAIR 2022, Shanghai, China, 21–23 October 2022*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 193–204.

104. Hasan, S.; Rivera, D.; Wu, X.; Durbin, E.; Christian, J.; Tourassi, G. Knowledge graph-enabled cancer data analytics. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 1952–1967. [CrossRef]

105. Deng, S.; Rangwala, H.; Ning, Y. Dynamic knowledge graph based multi-event forecasting. In Proceedings of the 26th ACM SIGKDD International Conference On Knowledge Discovery & Data Mining, Virtual, 6–10 July 2020; pp. 1585–1595.

106. Ilievski, F.; Garijo, D.; Chalupsky, H.; Divvala, N.; Yao, Y.; Rogers, C.; Li, R.; Liu, J.; Singh, A.; Schwabe, D.; et al. KGTK: A toolkit for large knowledge graph manipulation and analysis. In *The Semantic Web-ISWC 2020, Proceedings of the 19th International Semantic Web Conference, Athens, Greece, 2–6 November 2020;* Springer: Berlin/Heidelberg, Germany, 2020; pp. 278–293.

107. Chami, I.; Wolf, A.; Juan, D.; Sala, F.; Ravi, S.; Ré; C. Low-dimensional hyperbolic knowledge graph embeddings. *arXiv* **2020**, arXiv:2005.00545.

108. Wang, Q.; Li, M.; Wang, X.; Parulian, N.; Han, G.; Ma, J.; Tu, J.; Lin, Y.; Zhang, H.; Liu, W.; et al. COVID-19 literature knowledge graph construction and drug repurposing report generation. *arXiv* **2020**, arXiv:2007.00576.

109. Abu-Salih, B.; Al-Tawil, M.; Aljarah, I.; Faris, H.; Wongthongtham, P.; Chan, K.; Beheshti, A. Relational learning analysis of social politics using knowledge graph embedding. *Data Min. Knowl. Discov.* **2021**, *35*, 1497–1536. [CrossRef]

110. Sosa, D.; Derry, A.; Guo, M.; Wei, E.; Brinton, C.; Altman, R. A literature-based knowledge graph embedding method for identifying drug repurposing opportunities in rare diseases. In *Pacific Symposium on Biocomputing 2020*; World Scientific: Singapore, 2019; pp. 463–474.

111. Zhu, Y.; Che, C.; Jin, B.; Zhang, N.; Su, C.; Wang, F. Knowledge-driven drug repurposing using a comprehensive drug knowledge graph. *Health Inform. J.* **2020**, *26*, 2737–2750. [CrossRef] [PubMed]

112. Shen, W.; Zeng, X.; Zhu, F.; Wang, Y.; Qin, C.; Tan, Y.; Jiang, Y.; Chen, Y. Out-of-the-box deep learning prediction of pharmaceutical properties by broadly learned knowledge-based molecular representations. *Nat. Mach. Intell.* **2021**, *3*, 334–343. [CrossRef]

113. Waagmeester, A.; Stupp, G.; Burgstaller-Muehlbacher, S.; Good, B.; Griffith, M.; Griffith, O.; Hanspers, K.; Hermjakob, H.; Hudson, T.; Hybiske, K.; et al. Wikidata as a knowledge graph for the life sciences. *Elife* **2020**, *9*, e52614. [CrossRef] [PubMed]

114. Mohamed, S.; Nounu, A.; Nováček, V. Biological applications of knowledge graph embedding models. *Briefings Bioinform.* **2021**, *22*, 1679–1693. [CrossRef] [PubMed]

115. Johnson, A.; Pollard, T.; Shen, L.; Lehman, L.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.; Celi, L.; Mark, R. MIMIC-III, a freely accessible critical care database. *Sci. Data* **2016**, *3*, 160035. [CrossRef] [PubMed]

116. Protégé. Available online: https://protege.stanford.edu/ (accessed on 10 May 2023).